Strong Local Optimality for a Bang-Bang-Singular Extremal: General Constraints

JOTA manuscript No.
(will be inserted by the editor)

Strong Local Optimality for a Bang-Bang-Singular Extremal: General Constraints

Laura Poggiolini · Gianna Stefani

Communicated by Aram Arutyunov

This is a post-peer-review, pre-copyedit version of an article published in JOTA. The final authenticated version is available online at: http://dx.doi.org/10.1007/s10957-020-01700-2

Abstract The paper provides second order sufficient conditions for the strong local optimality of bang-bang-singular extremals in a Mayer problem with general end point constraints. The sufficient conditions are expressed as a strengthening of the necessary ones plus the coerciveness of a suitable quadratic form related to a sub-problem of the given one. The sufficiency of the given conditions is proven via Hamiltonian methods.

Keywords sufficient conditions, singular control, second variation, Hamiltonian methods.

Mathematics Subject Classification (2000) 49K15 · 49J15 · 93C10

1 Introduction

The aim of this paper is twofold: on one hand we extend the results obtained in [1], on the other hand we show how the fixed-free case studied there is a case study in the Hamiltonian approach. The results were announced in [2].

In [1] the authors gave sufficient second order conditions for the strong local optimality of a reference extremal trajectory in a Mayer problem with fixed-free end points constraints. Here, we consider the case when the trajectory has the same control structure, but the end points constraints are smooth submanifolds of the state space, thus allowing one also for abnormal extremals.

Laura Poggiolini, Corresponding author
Dipartimento di Matematica e Informatica "Ulisse Dini",
Università degli Studi di Firenze, Firenze, Italy
laura.poggiolini@unifi.it

Gianna Stefani,
Dipartimento di Matematica e Informatica "Ulisse Dini",
Università degli Studi di Firenze, Firenze, Italy
gianna.stefani@unifi.it

We show that, under our sufficient conditions, the extremal trajectory is a strict strong local minimizer (Definition 1) in the normal case and it is an isolated admissible trajectory in the abnormal one. The sufficient conditions are those, which allow us to apply the Hamiltonian approach described in Section 4. In particular, they include:

– Regularity assumptions which permit to define an overmaximized Hamiltonian; see Section 3.2.
– Coerciveness of a suitable second variation; see Section 5, which allows us to have an invertible projected overmaximized flow.

For a general introduction to the Hamiltonian approach to strong local optimality, the reader is referred to [3]. We also recall that sufficient conditions have been proven - via Hamiltonian methods - in several cases, whose common feature is that the dynamics is control affine.

We recall the papers [4, 5], which deal with bang-bang extremals with simple switches and [6,7], where the authors consider bang-bang extremals having a double switch. In [8], the author considers a singular extremal in a Mayer problem and introduces the overmaximized Hamiltonian. Indeed, the fact that the extremal is singular makes it impossible to exploit the Hamiltonian approach via the maximized Hamiltonian function; see [9, Section 3.1]. The results were extended to the singular multi-input case in [10,11]. The paper [9] faced the problem of concatenations of bang and singular arcs in the minimum time problem. Sufficient conditions have also been given in the case where the control takes values in a closed ball; see [12–15].

Sufficient conditions to various kinds of local optimality have been provided via other methods by several authors. We would like to mention the book [16] and the references therein. Important results have been obtained in [17] for bang-singular concatenations, while [18] provides some numerical applications. Finally, [19] provides sufficient second order conditions for optimal control problems on time scales. We conclude by mentioning [16, 20, 26], which motivated us to study the bang-bang-singular case.

2 Statement of the Problem and Notation

2.1 The Problem

Let N_0, N_T be smooth manifolds in $\mathbb{R}^n$, let $X_1, \ldots, X_m$ be distinct smooth vector fields and let $\mathcal{X}(x)$ be their convex hull at each point $x \in \mathbb{R}^n$. By smooth we mean C^∞, although C^2-regularity both on the constraint manifolds and on the vector fields would suffice.

For a precise definition of bang and totally or partially singular trajectories, the reader is referred to [1, Definition 1.2], where also an equivalent formulation as a multi-input optimal control problem is given. Here, we consider the following Mayer problem on a fixed time interval $[0, T]$:

$$\text{minimize} \quad C(\xi) = c_0(\xi(0)) + c_T(\xi(T)) \text{ subject to} \tag{1a}$$

$$\dot{\xi}(t) \in \mathcal{X}(\xi(t)), \quad \xi(0) \in N_0, \quad \xi(T) \in N_T, \tag{1b}$$

where solutions of (1b) are meant in the Carathéodory sense. We assume there exists a bang-bang-singular reference trajectory $\widehat{\xi}$, whose optimality we want to investigate. More precisely we aim at giving sufficient conditions for the *strong local optimality* of $\widehat{\xi}$ according to the following definition:

Definition 1 An admissible trajectory $\widehat{\xi}\colon [0,T] \to \mathbb{R}^n$ is a *strong local minimizer* of problem (1a)–(1b) if it is a minimizer among the admissible trajectories which are close to $\widehat{\xi}$ in the C^0 topology i.e. there exists $\varepsilon > 0$ such that $C(\xi) \leq C(\widehat{\xi})$ for any adissible trajectory ξ for problem (1b) satisfying $\left\| \xi - \widehat{\xi} \right\|_{C^0} := \max_{t\in[0,T]} \left\| \xi(t) - \widehat{\xi}(t) \right\| < \varepsilon$. If $C(\widehat{\xi}) < C(\xi)$ for any $\xi \neq \widehat{\xi}$ we say that $\widehat{\xi}$ is a *strict* strong local minimizer.

The trajectory $\widehat{\xi}$ has the following structure: there exist times $\widehat{\tau}_1, \widehat{\tau}_2$, called *switching times*, $0 < \widehat{\tau}_1 < \widehat{\tau}_2 < T$, vector fields h_1, h_2, $h_3 \in \{X_1, \dots X_m\}$, $h_2 \neq h_3$, and a function $\widehat{v} \in C^0([\widehat{\tau}_2, T])$ taking values in the open interval $]0,1[$ such that $\widehat{\xi}$ is a solution to

$$
\begin{aligned}
&\dot{\xi}(t) = h_1(\xi(t)) && t \in]0, \widehat{\tau}_1[, \\
&\dot{\xi}(t) = h_2(\xi(t)) && t \in]\widehat{\tau}_1, \widehat{\tau}_2[, \\
&\dot{\xi}(t) = \widehat{v}(t) h_3(\xi(t)) + (1 - \widehat{v}(t))\, h_2(\xi(t)) && t \in]\widehat{\tau}_2, T[, \\
&\xi(0) \in N_0, \quad \xi(T) \in N_T.
\end{aligned}
$$

Setting $f_{\mathrm{d}} := h_3 - h_2$, we can define the time-dependent reference vector field

$$
\widehat{f}_t := \begin{cases} h_1 & t \in [0, \widehat{\tau}_1[, \\ h_2 & t \in [\widehat{\tau}_1, \widehat{\tau}_2[, \\ h_2 + \widehat{v}(t) f_{\mathrm{d}} & t \in [\widehat{\tau}_2, T]. \end{cases}
$$

Notice that the continuity assumption on $\widehat{v}$ is not restrictive, in view of the SGLC condition which we are going to assume; see Assumption 4.

2.2 Notation

For the sake of conciseness and clarity we use notation from differential geometry. Indeed, this notation helps to realize that the discussion is intrinsic and the result can be applied to problems on manifolds. In particular we denote the trivial cotangent bundle $(\mathbb{R}^n)^* \times \mathbb{R}^n$ as $T^*\mathbb{R}^n$ and we distinguish between points in $\mathbb{R}^n$, usually denoted as x, and tangent vectors to $\mathbb{R}^n$, denoted as δx.

If N is a smooth manifold we denote as T_xN and T_x^*N the tangent and cotangent spaces to N at a point $x \in N$, respectively. The symbol $T_x^{\perp}N$ denotes the linear subspace of one-forms in $(\mathbb{R}^n)^*$ which are null on T_xN.

Given a C^1 vector field f on N, we denote as $\exp tf(x)$ the flow at time t emanating from a point x at time 0, i.e. $\exp tf(x)$ is the solution to

$$
\dot{\xi}(t) = f(\xi(t)), \quad \xi(0) = x.
$$

If g is another C^1 vector field, then the Lie bracket between f and g is denoted as $[f,g]$, i.e. $[f,g](x) := \mathrm{D}g(x)f(x) - \mathrm{D}f(x)g(x)$.

If $a\colon N \to \mathbb{R}$ is a C^2 function, $\mathrm{d}a(x)$ is its differential at $x \in N$. If, at some point $\overline{x}$ one has $\mathrm{d}a(\overline{x}) = 0$, then the second derivative of a at $\overline{x}$ is a well defined bilinear form on $T_{\overline{x}}N$ which we denote as $\mathrm{D}^2 a(\overline{x})$. $L_f a\,(x) := \langle \mathrm{d}a(x)\,,\, f(x)\rangle$ is the Lie derivative of a with respect to the vector field f at the point x. The symbol $\mathrm{d}a_*\delta x$ denotes the couple $\big(\mathrm{D}^2 a(x)(\delta x, \cdot), \delta x\big) \in T_x^*\mathbb{R}^n \times T_x\mathbb{R}^n$ whenever the point x is clear from the context.

If G is a C^1 map from a manifold X to a manifold Y, its tangent map at a point $x \in X$ is denoted as T_xG, or just as G_* if x is clear from the context.

We also use some basic element of the theory of symplectic manifolds referred to the trivial cotangent bundle $T^*\mathbb{R}^n$; see e.g. [3].

Let $\pi\colon \ell = (p, x) \in T^*\mathbb{R}^n \mapsto x \in \mathbb{R}^n$ be the canonical projection, and denote by $\boldsymbol{s} := \sum_{i=1}^n p^i\,\mathrm{d}x_i$ the canonical Liouville one–form on $T^*\mathbb{R}^n$. Given a time-dependent smooth Hamiltonian $H_t\colon T^*\mathbb{R}^n \to \mathbb{R}$ the canonical symplectic two-form $\boldsymbol{\sigma} = \mathrm{d}\boldsymbol{s} = \sum_{i=1}^n \mathrm{d}p^i \wedge \mathrm{d}x_i$ allows one to define a Hamiltonian vector field $\overrightarrow{H}_t$. Indeed one sets $\boldsymbol{\sigma}\left(V, \overrightarrow{H}_t(\ell)\right) = \langle \mathrm{d}H_t(\ell)\,,\, V\rangle$, for any $V \in T_\ell T^*\mathbb{R}^n$, i.e. $\overrightarrow{H}_t(\ell) = (-\,\partial_x H_t(\ell), \partial_p H_t(\ell))$, for any $\ell = (p, x) \in T^*\mathbb{R}^n$.

We recall that any vector field f on $\mathbb{R}^n$ defines a Hamiltonian

$$F\colon \ell = (p, x) \in T^*\mathbb{R}^n \mapsto \langle p\,,\, f(x)\rangle \in \mathbb{R}.$$

In particular we denote by H_1, H_2, H_3 and F_d the Hamiltonians associated with h_1, h_2, h_3 and f_d, respectively. Moreover H_{12} and H_{23} denote the Hamiltonians associated with the brackets $[h_1, h_2]$ and $[h_2, h_3]$, respectively. Analogously H_{223} and H_{323} are the Hamiltonians associated with $[h_2, [h_2, h_3]]$ and $[h_3, [h_2, h_3]]$, respectively. Finally $\widehat{F}_t$ denotes the time-dependent Hamiltonian function obtained by lifting the reference vector field $\widehat{f}_t$ and $\widehat{\mathcal{F}}_t$ denotes its flow from time T, i.e. the flow associated with its Hamiltonian vector field. Indeed, due to the construction required to carry on our proof, we deal with flows starting from time T and evolving backwards in time. In particular the flow $\widehat{S}_t$ of $\widehat{f}_t$ is a local diffeomorphism defined in a neighbourhood of $\widehat{\xi}(T)$.

We conclude the section introducing the maximized Hamiltonian $H^{\max}$:

$$H^{\max}(\ell) := \max\{\langle p\,,\, Y\rangle \colon Y \in \mathcal{X}(x)\} \qquad \forall \ell = (p, x) \in T^*\mathbb{R}^n.$$

3 Regularity Assumptions and Consequences

The first assumption is the main necessary condition for optimality, i.e. Pontryagin Maximum Principle (PMP):

Assumption 1 (PMP) *There exist $p_0 \in \{0, 1\}$ and an absolutely continuous mapping $\widehat{\mu}\colon [0, T] \to (\mathbb{R}^n)^*$, $p_0 + |\widehat{\mu}(T)| \neq 0$, such that a.e. $t \in [0, T]$*

$$\begin{aligned}
&\dot{\widehat{\mu}}(t) = -\widehat{\mu}(t)\,\mathrm{D}\widehat{f}_t(\widehat{\xi}(t)),\\
&\widehat{\mu}(0) \in p_0\,\mathrm{d}c_0(\widehat{\xi}(0)) + T^{\perp}_{\widehat{\xi}(0)}N_0, \quad \widehat{\mu}(T) \in -p_0\,\mathrm{d}c_T(\widehat{\xi}(T)) + T^{\perp}_{\widehat{\xi}(T)}N_T,\\
&\widehat{F}_t(\widehat{\mu}(t), \widehat{\xi}(t)) = H^{\max}(\widehat{\mu}(t), (\widehat{\xi}(t)).
\end{aligned}$$

$\widehat{\mu}$ is called adjoint covector and $\widehat{\xi}$ is called a state extremal *while the couple $\widehat{\lambda}(t) := \big(\widehat{\mu}(t), \widehat{\xi}(t)\big) \in T^*\mathbb{R}^n$ is called an* extremal.

We denote the terminal and switching points of $\widehat{\lambda}$ and $\widehat{\xi}$ as follows:

$$\begin{array}{llll}
\widehat{\ell}_0 := \widehat{\lambda}(0), & \widehat{\ell}_1 := \widehat{\lambda}(\widehat{\tau}_1), & \widehat{\ell}_2 := \widehat{\lambda}(\widehat{\tau}_2), & \widehat{\ell}_T := \widehat{\lambda}(T),\\
\widehat{x}_0 := \widehat{\xi}(0) = \pi\widehat{\ell}_0, & \widehat{x}_1 := \widehat{\xi}(\widehat{\tau}_1) = \pi\widehat{\ell}_1, & \widehat{x}_2 := \widehat{\xi}(\widehat{\tau}_2) = \pi\widehat{\ell}_2, & \widehat{x}_T := \widehat{\xi}(T) = \pi\widehat{\ell}_T.
\end{array}$$

3.1 Regularity Conditions

In this section we state the regularity conditions, i.e. we require strict inequalities whenever necessary conditions yield mild ones; see [1, Section 2.2].

Along each bang arc we assume that only the reference vector field gives the maximized Hamiltonian.

Assumption 2 (Regularity along the bang arcs)

$$\begin{aligned} H_1(\widehat{\lambda}(t)) &> \langle \widehat{\mu}(t)\,,\, Y\rangle \qquad \forall Y \in \mathcal{X}(\widehat{\xi}(t)) \setminus \{h_1(\widehat{\xi}(t))\}, \quad \forall t \in [0, \widehat{\tau}_1[, \\ H_2(\widehat{\lambda}(t)) &> \langle \widehat{\mu}(t)\,,\, Y\rangle \qquad \forall Y \in \mathcal{X}(\widehat{\xi}(t)) \setminus \{h_2(\widehat{\xi}(t))\}, \quad \forall t \in]\widehat{\tau}_1, \widehat{\tau}_2[. \end{aligned}$$

Along the singular arc we assume that only the vector fields in the edge defined by h_2 and h_3 give the maximized Hamiltonian.

Assumption 3 (Regularity along the singular arc) *For any $a \in [0, 1]$ and any $t \in [\widehat{\tau}_2, T]$*

$$H_2(\widehat{\lambda}(t)) + \widehat{v}(t)F_{\mathrm{d}}(\widehat{\lambda}(t)) > \langle \widehat{\mu}(t)\,,\, Y\rangle \qquad \forall Y \in \mathcal{X}(\widehat{\xi}(t)),\ Y \neq (h_2 + af_{\mathrm{d}})\,(\widehat{\xi}(t)).$$

Remark 1 The above regularity assumptions imply that $h_1(\widehat{\xi}(t))$ and $h_2(\widehat{\xi}(t))$ are vertexes of $\mathcal{X}(\widehat{\xi}(t))$ for any $t \in (0, \widehat{\tau}_1)$ and $t \in (\widehat{\tau}_1, \widehat{\tau}_2)$, respectively and that both $h_2(\widehat{\xi}(t))$ and $h_3(\widehat{\xi}(t))$ are vertexes of $\mathcal{X}(\widehat{\xi}(t))$ for any $t \in [\widehat{\tau}_2, T]$.

We assume the *strong generalised Legendre condition*. Namely, setting

$$\mathbb{L}(\ell) := (H_{323} - H_{223})(\ell) = \langle p\,,\, [f_{\mathrm{d}}, [h_2, f_{\mathrm{d}}]]\,(x)\rangle, \qquad \ell = (p, x) \in T^*\mathbb{R}^n,$$

we assume the following.

Assumption 4 (SGLC) *For all $t \in [\widehat{\tau}_2, T]$ there holds $R(t) := \mathbb{L}(\widehat{\lambda}(t)) > 0$.*

Remark 2 In [1] the authors prove that whenever SGLC holds, then the singular control is smooth on $[\widehat{\tau}_2, T]$.

Assumption 5 (Regularity at the switching times)

$$H_{12}(\widehat{\ell}_1) = \langle \widehat{\mu}(\widehat{\tau}_1)\,,\, [h_1, h_2]\,(\widehat{x}_1)\rangle > 0, \tag{2}$$

$$H_{223}(\widehat{\ell}_2) = \langle \widehat{\mu}(\widehat{\tau}_2)\,,\, [h_2, [h_2, f_{\mathrm{d}}]]\,(\widehat{x}_2)\rangle < 0. \tag{3}$$

Remark 3 Inequality (2) is the classical regularity condition for bang-bang concatenations, see [21], while inequality (3) implies that the reference vector field is discontinuous at $\widehat{\tau}_2$. By the necessary conditions, both $F_{\mathrm{d}} \circ \widehat{\lambda}(t)$ and $\dfrac{\mathrm{d}}{\mathrm{d}t} F_{\mathrm{d}} \circ \widehat{\lambda}(t)$ are null on the interval $[\widehat{\tau}_2, T]$, so that $\lim\limits_{t \to \widehat{\tau}_2^+} \dfrac{\mathrm{d}^k}{\mathrm{d}t^k} F_{\mathrm{d}} \circ \widehat{\lambda}(t) = 0$ for any $k \in \mathbb{N}$. Inequality (3) means that $\lim\limits_{t \to \widehat{\tau}_2^-} \dfrac{\mathrm{d}^2}{\mathrm{d}t^2} F_{\mathrm{d}} \circ \widehat{\lambda}(t) \neq 0$.

3.2 Geometry Near the Reference Extremal

It is easy to prove that SGLC implies that there exists a neighbourhood $\mathcal{O}_{\mathrm{s}}$ of the range of the singular arc $\widehat{\lambda}([\widehat{\tau}_2, T])$ in $T^*\mathbb{R}^n$ such that the sets

$$\Sigma := \{\ell \in \mathcal{O}_{\mathrm{s}} \colon F_{\mathrm{d}}(\ell) = 0\}\,, \qquad \mathcal{S} := \{\ell \in \Sigma \colon H_{23}(\ell) = 0\}$$

are smooth simply connected manifolds of codimension 1 and 2, respectively. More precisely $\overrightarrow{H}_{23}$ is transverse to Σ in $\mathcal{O}_{\mathrm{s}}$, while $\overrightarrow{F}_{\mathrm{d}}$ is tangent to Σ and transverse to $\mathcal{S}$ in Σ; see [9].

In [1] the authors prove that under Assumptions 1-5, there exists a tubular neighbourhood $\mathcal{O}$ of the graph of $\widehat{\lambda}$ and a piecewise C^1 Hamiltonian function

$$H \colon (t, \ell) \in \mathcal{O} \mapsto H_t(\ell) \in \mathbb{R}$$

that satisfies the regularity and overmaximization properties required to pursue the Hamiltonian approach to strong local optimality described in [22]. The construction of H_t is based on the techniques developed in [8] for singular extremals and then extended in [9] for concatenations of bang and singular arcs. We briefly sketch the construction; see [1, 9] for details.

Proposition 1 *Possibly restricting the neighbourhood $\mathcal{O}_{\mathrm{s}}$, there exists a smooth Hamiltonian function, $\widetilde{H}_2 \colon \mathcal{O}_{\mathrm{s}} \to \mathbb{R}$, satisfying the following properties*

- $\widetilde{H}_2(\ell) \geq H_2(\ell)$ *for any* $\ell \in \Sigma$*. Equality holds if and only if* $\ell \in \mathcal{S}$*.*
- $\overrightarrow{\widetilde{H}}_2$ *is tangent to* Σ *and* $\overrightarrow{\widetilde{H}}_2(\ell) = \overrightarrow{H}_2(\ell)$ *for any* $\ell \in \mathcal{S}$*.*

By the regularity assumptions there exists $\varepsilon > 0$ such that the following time-dependent Hamiltonian is well-defined in a neighbourhood of the graph of the reference extremal $\widehat{\lambda}$,

$$H_t(\ell) := \begin{cases} \widetilde{H}_2(\ell) + \widehat{v}(t) F_{\mathrm{d}}(\ell), & t \in [\widehat{\tau}_2, T], \\ \widetilde{H}_2(\ell), & H_{23}(\ell) > 0, \ t \in [\widehat{\tau}_2 - \varepsilon, \widehat{\tau}_2[, \\ H^{\max}(\ell), & \text{else.} \end{cases}$$

The flow associated with this Hamiltonian and starting from time T is denoted as $\mathcal{H}_t$ and it is well defined in a neighbourhood of $\widehat{\ell}_T$ as described in the following propositions.

Proposition 2 *There exists a neighbourhood $\mathcal{O}_2$ of $\widehat{\ell}_2$ such that the followings hold.*

- *By means of the implicit function theorem the equation*

$$H_{23} \circ \exp(t_2 - \widehat{\tau}_2)\overrightarrow{\widetilde{H}}_2(\ell) = 0$$

defines a smooth function $t_2 \colon \mathcal{O}_2 \to \mathbb{R}$ such that $t_2(\widehat{\ell}_2) = \widehat{\tau}_2$. We set $\tau_2(\ell) := \min\{t_2(\ell), \widehat{\tau}_2\}$.

- *By means of the implicit function theorem the equation*

$$(H_1 - H_2) \circ \exp(\tau_1 - \tau_2(\ell))\overrightarrow{H}_2 \circ \exp(\tau_2(\ell) - \widehat{\tau}_2)\overrightarrow{\widetilde{H}}_2(\ell) = 0$$

defines a piecewise C^1 function $\tau_1 \colon \mathcal{O}_2 \to \mathbb{R}$ such that $\tau_1(\widehat{\ell}_2) = \widehat{\tau}_1$. Moreover, for any $\delta\ell \in T_{\widehat{\ell}_2} T^*\mathbb{R}^n$ the differential of τ_1 at $\widehat{\ell}_2$ is given by

$$\langle \mathrm{d}\tau_1(\widehat{\ell}_2)\,,\, \delta\ell \rangle = \frac{\boldsymbol{\sigma}\big(\exp(\widehat{\tau}_1 - \widehat{\tau}_2)\overrightarrow{H}_{2*}\delta\ell, (\overrightarrow{H}_1 - \overrightarrow{H}_2)(\widehat{\ell}_1)\big)}{H_{12}(\widehat{\ell}_1)}. \tag{4}$$

Proposition 3 *There exists a neighbourhood $\mathcal{O}_T$ of $\widehat{\ell}_T$ in Σ such that the flow $\mathcal{H}_t$ emanating from $\mathcal{O}_T$ is well defined, Lipschitz continuous and piecewise C^1. Moreover $\widehat{\lambda}(t) = \mathcal{H}_t(\widehat{\ell}_T)$. Let $\widetilde{\ell} := \mathcal{H}_{\widehat{\tau}_2}(\ell)$, then*

- *$\mathcal{H}_t(\ell) \in \Sigma$ for $(t, \ell) \in [\tau_2(\widetilde{\ell}), T] \times \mathcal{O}_T$.*
- *$\mathcal{H}_t$ is an over-maximized flow, i.e.*
  - *$H_t(\mathcal{H}_t(\ell)) \geq H^{\max}(\mathcal{H}_t(\ell))$ for any $(t, \ell) \in [\tau_2(\widetilde{\ell}), T] \times \mathcal{O}_T$ and equality holds if and only if $\mathcal{H}_t(\ell) \in \mathcal{S}$.*
  - *$H_t(\mathcal{H}_t(\ell)) = H^{\max}(\mathcal{H}_t(\ell))$ for any $t \in [0, \tau_2(\widetilde{\ell})]$.*
- *The linearised flow at time $t = 0$ is given by*

$$\mathcal{H}_{0*}\delta\ell = \widehat{\mathcal{F}}_{0*}\left(-\langle \mathrm{d}\tau_1(\widehat{\ell}_2)\,,\, \widetilde{\delta\ell}\rangle \overrightarrow{K}(\widehat{\ell}_T) + \widehat{\mathcal{F}}^{-1}_{\widehat{\tau}_2 *}\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell\right). \tag{5}$$

4 The Hamiltonian Approach

The leading idea of the Hamiltonian approach is to compare the costs of neighbouring admissible trajectories by lifting them to the cotangent bundle; see [22] for a detailed explanation. This is done via the Hamiltonian flow $\mathcal{H}_t$ introduced in the previous section and it is possible thanks to the coerciveness of the extended second variation defined in Section 5.

More precisely, we consider the flow emanating from a horizontal Lagrangian manifold Λ such that $\widehat{\ell}_T \in \Lambda \subset \Sigma$. Its construction is granted by the reduction of the problem to a free final point one, see Section 5.3, while the existence of the lifting is proven studying the coerciveness of the extended second variation of the fixed-free problem studied in [1].

The lifting allows one to estimate the difference of the cost of admissible trajectories by the variation of a function of their initial points, i.e. the optimal control problem is transformed into a finite dimensional optimisation problem.

4.1 A Hamiltonian Sufficient Condition

In this section, we prove an extension of [1, Theorem 4.2], which takes into account the cost on the initial point. Namely, let $\widetilde{c}$ be a smooth function defined on a neighbourhood of $\widehat{x}_T$ such that

$$L_{f_{\mathrm{d}}}\widetilde{c} \equiv 0, \qquad \mathrm{d}(-\widetilde{c})(\widehat{x}_T) = \widehat{\mu}(T). \tag{6}$$

We consider the problem with free final point given by

$$\text{minimize } \ \widetilde{C}(\xi) := \alpha(\xi(0)) + \widetilde{c}(\xi(T)) \text{ subject to} \tag{7a}$$

$$\dot{\xi}(t) \in \mathcal{X}(\xi(t)), \quad \text{a.e. } t \in [0, T], \quad \xi(0) \in N_0. \tag{7b}$$

Theorem 1 *Let Λ be the horizontal Lagrangian submanifold defined by*

$$\Lambda := \{(\mathrm{d}(-\widetilde{c})(x), x)\,,\ x \in \pi(\mathcal{O}_T)\}\,.$$

Assume the followings hold:

1. *There exists a neighbourhood $\mathcal{U}$ of the graph of $\widehat{\xi}$ such that the map*

$$\mathrm{id} \times \pi\mathcal{H} \colon (t, \ell) \in [0,T] \times \Lambda \mapsto (t, \pi\mathcal{H}_t(\ell)) \in \mathcal{U} \subset [0,T] \times \mathbb{R}^n \tag{8}$$

is locally Lipschitz invertible.

2. $\boldsymbol{\sigma}\left(\mathrm{d}\alpha_*(\pi\mathcal{H}_0)_*\delta\ell, \mathcal{H}_{0*}\delta\ell\right) > 0$ *for any* $\delta\ell \in T_{\widehat{\ell}_T}\Lambda$ *such that* $(\pi\mathcal{H}_0)_*\delta\ell$ *is in* $T_{\widehat{x}_0}N_0 \setminus \{0\}$.

Then $\widehat{\xi}$ is a strict strong locally optimal trajectory for problem (7a)–(7b).

Proof By (6), $\Lambda \subset \Sigma$ so that the flow in (8) satisfies Proposition 3. Clearly $(\mathrm{id} \times \pi\mathcal{H})^{-1}(t, \widehat{\xi}(t)) = (t, \widehat{\ell}_T)$ for any $t \in [0,T]$. Let $\xi \colon [0,T] \to \mathbb{R}^n$ be an admissible trajectory for (7b) whose graph is in $\mathcal{U}$ and let

$$(t, \ell(t)) := (\mathrm{id} \times \pi\mathcal{H})^{-1}(t, \xi(t)), \quad \lambda(t) := \mathcal{H}_t(\ell(t)) = (\mu(t), \xi(t)), \quad t \in [0,T].$$

Let $\psi_0 \colon [0,1] \to N_0$ be a curve such that $\psi_0(0) = \widehat{x}_0$, $\psi_0(1) = \xi(0)$ and define $\varphi_0 := (\pi\mathcal{H}_0)^{-1} \circ \psi_0 \colon [0,1] \to \Lambda$. Also let $\varphi_T \colon [0,1] \to \Lambda$ be a curve such that $\varphi_T(0) = \ell(T)$, $\varphi_T(1) = \widehat{\ell}_T$. We can consider the closed path in $[0,T] \times \Lambda$ obtained by the concatenation of the curves

$$t \mapsto (t, \ell(t)), \qquad s \mapsto (T, \varphi_T(s)), \qquad t \mapsto (T - t, \widehat{\ell}_T), \qquad s \mapsto (0, \varphi_0(s)).$$

The one-form $\boldsymbol{\omega} := \mathcal{H}^*\left(\boldsymbol{s} - H_t\,\mathrm{d}t\right)$ is exact on $[0,T] \times \Lambda$; see [22]. Thus there exists a Lipschitz continuous function $\theta \colon (t, \ell) \in [0,T] \times \Lambda \mapsto \theta_t(\ell) \in \mathbb{R}$ such that $\mathrm{d}\theta = \omega$. In particular we can choose

$$\theta(t, \ell) = -\widetilde{c}(\pi\ell) - \int_t^T \left(\langle \mathcal{H}_s(\ell), \pi_*\overrightarrow{H}_s \circ \mathcal{H}_s(\ell)\rangle - H_s \circ \mathcal{H}_s(\ell)\right)\mathrm{d}s.$$

Integrating the one-form $\boldsymbol{\omega}$, we obtain

$$\begin{aligned} 0 = \oint \boldsymbol{\omega} = \int_{\mathrm{id}\times\ell} \left(\langle \mu(t), \dot{\xi}(t)\rangle - H_t(\lambda(t))\right)\mathrm{d}t + \int_{\varphi_T} \boldsymbol{s} \, + \\ - \int_{\mathrm{id}\times\widehat{\ell}_T} \left(\langle \widehat{\mu}(t), \dot{\widehat{\xi}}(t)\rangle - H_t(\widehat{\lambda}(t))\right)\mathrm{d}t + \int_{\varphi_0} \mathcal{H}_0^*\boldsymbol{s}. \end{aligned} \tag{9}$$

Since $\mathrm{d}\left(\theta_t \circ (\pi\mathcal{H}_t)^{-1}\right)(x) = \mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1}(x)$, see [22, Lemma 3.4], we get

$$\int_{\varphi_0} \mathcal{H}_0^*\boldsymbol{s} = \int_{\mathcal{H}_0\circ\varphi_0} \boldsymbol{s} = \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)(\xi(0)) - \theta_0(\widehat{\ell}_T).$$

Substituting in (9) we obtain

$$0 \le -\widetilde{c}(\widehat{x}_T) + \widetilde{c}(\xi(T)) + \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)(\xi(0)) - \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)(\widehat{x}_0),$$

hence

$$\widetilde{C}(\xi) - \widetilde{C}(\widehat{\xi}) \ge \left(\alpha - \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)\right)(\xi(0)) - \left(\alpha - \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)\right)(\widehat{x}_0). \tag{10}$$

By PMP and the properties of θ,

$$\mathrm{d}\left(\alpha - \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)\right)(\widehat{x}_0) = 0. \tag{11}$$

For any $\delta y_0 \in T_{\widehat{x}_0} N_0 \setminus \{0\}$ set $\delta\ell := (\pi\mathcal{H}_0)_*^{-1}\delta y_0$. By *2.*, we have

$$\mathrm{D}^2 \left(\alpha - \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right)\right)(\widehat{x}_0)[\delta y_0]^2 = \boldsymbol{\sigma}\left(\mathrm{d}\alpha_*\delta y_0, \mathcal{H}_{0*}\delta\ell\right) > 0. \tag{12}$$

Thus $\widehat{x}_0$ is a strict local minimum for $\alpha - \left(\theta_0 \circ (\pi\mathcal{H}_0)^{-1}\right) : N_0 \to \mathbb{R}$. Hence, by (10), $\widehat{\xi}$ is a strong local minimizer for the cost $\widetilde{C}$ constrained to (7b).

If $\widetilde{C}(\xi) = \widetilde{C}(\widehat{\xi})$, then $\xi(0) = \widehat{x}_0$ by (11)–(12). Proceeding as in the proof of Theorem 4.2 in [1] we obtain that $\widehat{\xi}$ is a strict strong local minimizer. □

5 The Extended Second Variation

The classical second variation of control affine systems is completely degenerate. We show here how to construct a non degenerate one.

We first extend to the whole $\mathbb{R}^n$ the cost functions p_0c_0 and p_0c_T in such a way that the transversality conditions hold on $\mathbb{R}^n$. Indeed, it is possible to find smooth functions α, $\beta\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\beta = p_0 c_T \quad \text{on } N_T, \qquad \widehat{\mu}(T) = \mathrm{d}(-\beta)(\widehat{x}_T) \quad \text{on } \mathbb{R}^n, \tag{13}$$
$$\alpha = p_0 c_0 \quad \text{on } N_0, \qquad \widehat{\mu}(0) = \mathrm{d}\alpha(\widehat{x}_0) \qquad \text{on } \mathbb{R}^n. \tag{14}$$

In the normal case ($p_0 = 1$) β and α are cost functions equivalent to the original ones while in the abnormal case ($p_0 = 0$) we study a problem with zero cost, thus saying that $\widehat{\xi}$ is a *strict* strong minimizer with $p_0 = 0$ means that it is isolated among admissible trajectories i.e. there exists $\varepsilon > 0$ such that there is no admissible trajectory ξ for problem (1b) satisfying $\left\|\xi - \widehat{\xi}\right\|_{C^0} < \varepsilon$.

We consider the problem of minimising the cost $\alpha(\xi(0)) + \beta(\xi(T))$ allowing only for perturbations of the reference control on the singular interval $]\widehat{\tau}_2, T[$ and for perturbations of the switching time $\widehat{\tau}_1$. A reparametrization of time on the interval $[0, \widehat{\tau}_2]$ allows to write the reduced problem as

$$\text{minimize } \alpha(\xi(0)) + \beta(\xi(T)) \quad \text{subject to}$$
$$\dot{\xi}(t) = \begin{cases} \upsilon_0(t)h_1(\xi(t)), & t \in]0, \widehat{\tau}_1[, \\ \upsilon_0(t)h_2(\xi(t)), & t \in]\widehat{\tau}_1, \widehat{\tau}_2[, \\ h_2(\xi(t)) + \upsilon(t)f_{\mathrm{d}}(\xi(t)), & t \in]\widehat{\tau}_2, T[, \end{cases}$$
$$\upsilon_0(t) > 0, \quad \int_0^{\widehat{\tau}_2} \upsilon_0(t)\,\mathrm{d}t = \widehat{\tau}_2, \quad \upsilon(t) \in (0,1), \quad \xi(0) \in N_0,\ \xi(T) \in N_T.$$

In order to obtain the second variation, we first pushforward the problem to the final time T, so that we obtain a problem evolving in a neighbourhood of $\widehat{x}_T$ and where the reference trajectory $\widehat{\xi}$ corresponds to the constant trajectory $t \mapsto \widehat{x}_T$. Indeed, setting $\delta := \min\{\min\{\widehat{\upsilon}(t),\ 1 - \widehat{\upsilon}(t)\} \colon t \in [\widehat{\tau}_2, T]\}$ and

$$\widehat{\alpha} := \alpha \circ \widehat{S}_0, \quad \widehat{N}_0 := \widehat{S}_0^{-1}(N_0), \quad g_t := \widehat{S}_{t*}^{-1} f_{\mathrm{d}} \circ \widehat{S}_t\,, \quad t \in [\widehat{\tau}_2, T],$$
$$k_i = \widehat{S}_{\widehat{\tau}_1 *}^{-1} h_i \circ \widehat{S}_{\widehat{\tau}_1}, \quad i = 1, 2, \qquad k := k_1 - k_2$$

we obtain the equivalent optimal control problem

$$
\begin{aligned}
&\text{minimize } \widehat{\alpha}(y)+\beta(x) \quad \text{subject to} \\
&\dot\eta(t) = \begin{cases} (\upsilon_0(t)-1)k_1(\eta(t)), & t \in]0,\widehat{\tau}_1[, \\ (\upsilon_0(t)-1)k_2(\eta(t)), & t \in]\widehat{\tau}_1,\widehat{\tau}_2[, \\ (\upsilon(t)-\widehat{\upsilon}(t))g_t(\eta(t)), & t \in]\widehat{\tau}_2, T[, \end{cases} \\
&\upsilon_0(t) > 0, \ \int_0^{\widehat{\tau}_2} (\upsilon_0(t)-1)\,\mathrm{d}t = 0, \ |\upsilon(t)-\widehat{\upsilon}(t)| < \delta, \\
&\eta(0) = y \in \widehat{N}_0, \quad \eta(T) = x \in N_T.
\end{aligned}
\tag{15}
$$

As in [1], we first write the second variation following [23]. Then, via a coordinate free version of Goh transformation, we extend it to a new quadratic form which we call *extended second variation*; see Appendix A for a detailed construction. Namely, consider the linear system on $[\widehat{\tau}_2, T]$

$$
\dot\zeta(t) = w(t)\dot{g}_t(\widehat{x}_T), \quad \zeta(\widehat{\tau}_2) = \delta y + \varepsilon_0\, k(\widehat{x}_T), \quad \zeta(T) = \delta x + \varepsilon_1 f_{\mathrm{d}}(\widehat{x}_T) \tag{16}
$$

and let $\delta e := (\delta x, \delta y, \varepsilon_0, \varepsilon_1, w) \in T_{\widehat{x}_T}N_T \times T_{\widehat{x}_T}\widehat{N}_0 \times \mathbb{R} \times \mathbb{R} \times L^2([\widehat{\tau}_2, T], \mathbb{R})$. We define the space of *extended admissible variations* as the Hilbert space

$$
\mathcal{W}_{\mathrm{ext}} := \{\delta e \colon \text{system (16) admits a solution}\}.
$$

The extended second variation of (15) is the quadratic form on $\mathcal{W}_{\mathrm{ext}}$ given by

$$
\begin{aligned}
J_{\mathrm{ext}}[\delta e]^2 = {} & \frac{1}{2}\,\mathrm{D}^2(\widehat{\alpha}+\beta)(\widehat{x}_T)[\delta y]^2 + \frac{\varepsilon_0^2}{2}\Big(L_k^2\beta\,(\widehat{x}_T) + L_{[k_2,k_1]}\beta\,(\widehat{x}_T)\Big) \\
& + \varepsilon_0 L_{\delta y} L_k \beta\,(\widehat{x}_T) - \frac{\varepsilon_1^2}{2} L_{f_{\mathrm{d}}}^2 \beta\,(\widehat{x}_T) - \varepsilon_1 L_{\delta x} L_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) \\
& + \frac{1}{2}\int_{\widehat{\tau}_2}^T \Big(2\,w(t) L_{\zeta(t)} L_{\dot{g}_t}\beta\,(\widehat{x}_T) + w(t)^2 R(t)\Big)\,\mathrm{d}t.
\end{aligned}
$$

Remark 4 J_{ext} is independent of the choice of α and β satisfying (13)–(14). Moreover $L_{[k_2,k_1]}\beta\,(\widehat{x}_T) = H_{12}(\widehat{\ell}_1)$ and $\dot{g}_t = \widehat{S}_{t*}^{-1}[h_2,h_3] \circ \widehat{S}_t \quad \forall t \in [\widehat{\tau}_2, T]$.

5.1 Coerciveness of the Extended Second Variation

In the forthcoming discussion, we study the coerciveness of the extended second variation, in the two following cases

1. $f_{\mathrm{d}}(\widehat{x}_T) \notin T_{\widehat{x}_T}N_T$. In this case we can choose β with properties (13) and such that $L_{f_{\mathrm{d}}}\beta \equiv 0$, so that $\{(\mathrm{d}(-\beta)(x), x)\} \subset \Sigma$. We define
$$
\widetilde{\beta} := \beta \quad \text{and} \quad V_T := T_{\widehat{x}_T}N_T \oplus \mathbb{R} f_{\mathrm{d}}(\widehat{x}_T).
$$
2. $f_{\mathrm{d}}(\widehat{x}_T) \in T_{\widehat{x}_T}N_T$. In this case $(-f_{\mathrm{d}}(\widehat{x}_T), 0, 0, 1, 0) \in \mathcal{W}_{\mathrm{ext}}$ and
$$
J_{\mathrm{ext}}[(-f_{\mathrm{d}}(\widehat{x}_T), 0, 0, 1, 0)]^2 = \frac{1}{2} L_{f_{\mathrm{d}}}^2 \beta\,(\widehat{x}_T)\,.
$$

Thus, a necessary condition for the coerciveness of J_{ext} is $L^2_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) > 0$. This condition does not depend on the choice of β and it implies that we are dealing with the normal case $p_0 = 1$.
Proceeding as in [1] consider the hyper-surface, locally defined near $\widehat{x}_T$,

$$\widetilde{M} := \left\{x \in \mathbb{R}^n \colon L_{f_{\mathrm{d}}}\beta\,(x) = 0\right\}$$

and the intersection $\widetilde{N}_T := \widetilde{M} \cap N_T$. Since $L^2_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) \neq 0$, $\widetilde{N}_T$ is a submanifold of N_T and its tangent space at $\widehat{x}_T$ is

$$T_{\widehat{x}_T}\widetilde{N}_T = \left\{\delta z \in T_{\widehat{x}_T}N_T \colon L_{\delta z}L_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) = 0\right\}.$$

We now extend $\beta|_{\widetilde{M}}$ as a constant function along the integral lines of f_{d}, i.e. for any $x = \exp(rf_{\mathrm{d}})(z)$, $z \in \widetilde{M}$ we set $\widetilde{\beta}(x) := \beta(z)$. In a sufficiently small neighbourhood $\mathcal{O}$ of $\widehat{x}_T$, the function $\widetilde{\beta}\colon \mathcal{O} \to \mathbb{R}$ is smooth. Moreover

$$\begin{aligned} &\widetilde{\beta}(\widehat{x}_T) = \beta(\widehat{x}_T), \quad && \mathrm{d}\widetilde{\beta}(\widehat{x}_T) = \mathrm{d}\beta(\widehat{x}_T) = -\widehat{\mu}(T), \\ &\widetilde{\beta}(x) \leq \beta(x), && L_{f_{\mathrm{d}}}\widetilde{\beta}\,(x) = 0 \quad \forall x \in \mathcal{O}. \end{aligned} \tag{17}$$

Finally we set $V_T := T_{\widehat{x}_T}N_T$.

We now show how, in both cases, the coerciveness of J_{ext} on $\mathcal{W}_{\mathrm{ext}}$ can be studied via the coerciveness of an equivalent quadratic form. Indeed, consider the boundary value problem

$$\dot{\zeta}(t) = w(t)\dot{g}_t(\widehat{x}_T), \quad \zeta(\widehat{\tau}_2) = \delta y + \varepsilon_0\, k(\widehat{x}_T), \quad \zeta(T) = \delta x, \tag{18}$$

$$\delta x \in \mathbb{R}^n, \quad \delta y \in \mathbb{R}^n, \quad \varepsilon_0 \in \mathbb{R}, \quad w \in L^2([\widehat{\tau}_2, T], \mathbb{R}), \tag{19}$$

and let $\mathcal{A} \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times L^2([\widehat{\tau}_2, T], \mathbb{R})$ be the Hilbert space of 4-uple $\delta e = (\delta x, \delta y, \varepsilon_0, w)$ such that (18)-(19) admits a solution.

Let the function $\widetilde{\beta}$ and the linear space V_T be defined according to the two cases above and set

$$\Gamma := \mathrm{D}^2\big(\widehat{\alpha} + \widetilde{\beta}\big)(\widehat{x}_T), \qquad J_0 := L^2_k\widetilde{\beta}\,(\widehat{x}_T) + H_{12}(\widehat{\ell}_1),$$

$$\begin{aligned} J[\delta e]^2 := {} & \frac{1}{2}\Gamma[\delta y]^2 + \frac{\varepsilon_0^2}{2}J_0 + \varepsilon_0 L_{\delta y}L_k\widetilde{\beta}\,(\widehat{x}_T) \\ & + \frac{1}{2}\int_{\widehat{\tau}_2}^{T}\Big(2\,w(t)L_{\zeta(t)}L_{\dot{g}_t}\widetilde{\beta}\,(\widehat{x}_T) + w(t)^2R(t)\Big)\,\mathrm{d}t. \end{aligned} \tag{20}$$

In both cases the coerciveness of J_{ext} on $\mathcal{W}_{\mathrm{ext}}$ implies the coerciveness of J on

$$\mathcal{W} := \left\{\delta e = (\delta x, \delta y, \varepsilon_0, w) \in \mathcal{A} \colon \delta x \in V_T,\ \delta y \in T_{\widehat{x}_T}\widehat{N}_0\right\}. \tag{21}$$

More precisely, in case 1. the coerciveness of J_{ext} on $\mathcal{W}_{\mathrm{ext}}$ is equivalent to the coerciveness of J on $\mathcal{W}$, while in case 2. the coerciveness of J_{ext} on $\mathcal{W}_{\mathrm{ext}}$ is equivalent to the coerciveness of J on $\mathcal{W}$ plus $L^2_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) > 0$; see Appendix B.

Remark 5 Notice that in both cases we have $f_{\mathrm{d}}(\widehat{x}_T) \neq 0$. Moreover, in the case $N_T = M$, the restriction of (20)–(21) to $\delta y = 0$ coincides with the extended second variation (16)-(17) in [1] where $\widetilde{\beta}$ is replaced by $\widetilde{c}$.

5.1.1 Exploiting the Coerciveness of J on $\mathcal{W}$

In order to successfully exploit the coerciveness of J, we introduce two subspaces of $\mathcal{W}$. Namely, let $\mathcal{V}_0 \subset \mathcal{W}_0 \subset \mathcal{W}$ be defined as:

$$\mathcal{V}_0 := \{\delta e \in \mathcal{W} \colon \delta y = 0, \varepsilon_0 = 0\}, \qquad \mathcal{W}_0 := \{\delta e \in \mathcal{W} \colon \delta y = 0\}.$$

Denote by $\mathcal{V}_0^{\perp_J}$ and $\mathcal{W}_0^{\perp_J}$ their orthogonal spaces in $\mathcal{W}$ with respect to the bilinear form associated with J. It is well known that the coerciveness of J on $\mathcal{W}$ is equivalent to the coerciveness of J on $\mathcal{V}_0$, on $\mathcal{W}_0 \cap \mathcal{V}_0^{\perp_J}$ and on $\mathcal{W} \cap \mathcal{W}_0^{\perp_J}$.

The main tools that we need are the linear Hamiltonian system and the bilinear form associated with J. Indeed we can define the minimized Hamiltonian associated with the linear-quadratic problem (18)–(20)

$$H''_t(\delta p, \delta x) := -\frac{1}{2R(t)} \Big(\langle \delta p \,,\, \dot g_t(\widehat{x}_T)\rangle + L_{\delta x} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T) \Big)^2 \tag{22}$$

and the Lagrangian subspace of the final transversality conditions

$$L''_T := \Big\{ (\delta p, \delta x) \colon \delta p \in V_T^{\perp},\ \delta x \in V_T \Big\}. \tag{23}$$

We consider the Hamiltonian system defined by (22)–(23)

$$\begin{cases} \dot\mu''(t) = \dfrac{1}{R(t)} \Big(\langle \mu''(t) \,,\, \dot g_t(\widehat{x}_T)\rangle + L_{\zeta''(t)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T) \Big) L_{(\cdot)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T)\,, \\ \dot\zeta''(t) = \dfrac{-1}{R(t)} \Big(\langle \mu''(t) \,,\, \dot g_t(\widehat{x}_T)\rangle + L_{\zeta''(t)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T) \Big) \dot g_t(\widehat{x}_T), \\ \big(\mu''(T), \zeta''(T)\big) = (\delta p, \delta x) \in L''_T, \end{cases} \tag{24}$$

and we denote its solution as

$$\mathcal{H}''_t \colon (\delta p, \delta x) \in L''_T \mapsto \big(\mu''_t(\delta p, \delta x), \zeta''_t(\delta p, \delta x)\big) \in (\mathbb{R}^n)^* \times \mathbb{R}^n.$$

We consider the bilinear form associated with J on $\mathcal{A}$

$$\begin{aligned} 2\, J(\delta e, \overline{\delta e}) &= \Gamma(\delta y, \overline{\delta y}) + \varepsilon_0 \overline{\varepsilon}_0 J_0 + \varepsilon_0 L_{\overline{\delta y}} L_k \widetilde{\beta}\,(\widehat{x}_T) + \overline{\varepsilon}_0 L_{\delta y} L_k \widetilde{\beta}\,(\widehat{x}_T) \\ &+ \int_{\widehat{\tau}_2}^T \Big(w(t) L_{\overline{\zeta}(t)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T) + \overline{w}(t) L_{\zeta(t)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T) + w(t)\overline{w}(t) R(t) \Big)\, \mathrm{d}t. \end{aligned}$$

If $p(t)$ is a solution of the differential equation

$$\dot p(t) = -w(t) L_{(\cdot)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T)\,,$$

we obtain

$$\begin{aligned} 2\, J(\delta e, \overline{\delta e}) &= \Gamma(\delta y, \overline{\delta y}) + \varepsilon_0 L_{\overline{\delta y}} L_k \widetilde{\beta}\,(\widehat{x}_T) + \langle p(\widehat{\tau}_2) \,,\, \overline{\delta y}\rangle \\ &+ \overline{\varepsilon}_0 \Big(\varepsilon_0 J_0 + L_{\delta y} L_k \widetilde{\beta}\,(\widehat{x}_T) + \langle p(\widehat{\tau}_2) \,,\, k(\widehat{x}_T)\rangle \Big) - \langle p(T) \,,\, \overline{\zeta}(T)\rangle \\ &+ \int_{\widehat{\tau}_2}^T \overline{w}(t) \Big(\langle p(t) \,,\, \dot g_t(\widehat{x}_T)\rangle + L_{\zeta(t)} L_{\dot g_t} \widetilde{\beta}\,(\widehat{x}_T) + w(t) R(t) \Big)\, \mathrm{d}t. \end{aligned} \tag{25}$$

Another fundamental tool relies on the property that the flow $\mathcal{H}''_t$, up to a linear transformation, is the differential of the flow $\mathcal{H}_t$. This relation can be obtained via the antisymplectic isomorphism

$$\iota\colon (\delta p, \delta x) \in (\mathbb{R}^n)^* \times \mathbb{R}^n \mapsto \delta\ell := \Big(-\delta p + \mathrm{D}^2\,(-\widetilde{\beta})(\widehat{x}_T)(\delta x, \cdot), \delta x\Big) \in T^*\mathbb{R}^n;$$

see subsequent property 3. Indeed

$$\iota \mathcal{H}''_t \iota^{-1} = \widehat{\mathcal{F}}^{-1}_{t*} \mathcal{H}_{t*} \qquad \forall t \in [\widehat{\tau}_2, T]. \tag{26}$$

Formula (26) was first proven in [23] where a symplectic isomorphism was considered. We recall the following properties of ι which we extensively use:

1. $\iota = \iota^{-1}$,
2. $L_T := \iota L''_T = \Big\{(\omega, \delta x) \in (\mathbb{R}^n)^* \times V_T : \omega|_{V_T} = \mathrm{D}^2(-\widetilde{\beta})(\widehat{x}_T)(\delta x, \cdot)\Big\}$,
3. $\boldsymbol{\sigma}\left(\iota(\delta p, \delta x), \iota(\overline{\delta p}, \overline{\delta x})\right) = \boldsymbol{\sigma}\left((\overline{\delta p}, \overline{\delta x}), (\delta p, \delta x)\right),\ \delta p, \overline{\delta p} \in (\mathbb{R}^n)^*,\ \delta x, \overline{\delta x} \in \mathbb{R}^n.$

Moreover, in order to describe the orthogonal subspaces introduced above in terms of the flow $\mathcal{H}_{t*}$ it is useful to introduce the vector field $\widetilde{k} := \widehat{S}_{\widehat{\tau}_2 *} k \circ \widehat{S}^{-1}_{\widehat{\tau}_2}$ and its associated Hamiltonian $\widetilde{K}(p, x) = \langle p\,,\, \widehat{S}_{\widehat{\tau}_2 *} k \circ \widehat{S}^{-1}_{\widehat{\tau}_2}(x)\rangle$.

5.1.2 The Coerciveness of J on $\mathcal{V}_0$

The coerciveness of J on $\mathcal{V}_0$ is characterized in the lemma below.

Lemma 1 *The following conditions are equivalent*

1. *The quadratic form J is coercive on $\mathcal{V}_0$.*
2. *If $\exists t \in [\widehat{\tau}_2, T]$ and $(\delta p, \delta x) \in L''_T$ such that $\pi_* \mathcal{H}''_t(\delta p, \delta x) = 0$, then*
$$\delta x = 0, \qquad \mathcal{H}''_s(\delta p, 0) = (\delta p, 0) \quad \forall s \in [t, T].$$
3. *If $\exists t \in [\widehat{\tau}_2, T]$ and $\delta\ell \in L_T$ such that $(\pi \mathcal{H}_t)_*(\delta\ell) = 0$, then*
$$\pi_* \delta\ell = 0, \qquad \mathcal{H}_s(\delta\ell) = \delta\ell \quad \forall s \in [t, T].$$

Proof For the equivalence of *1.* and *2.*; see [24]. Take advantage of (26) to obtain the equivalence of *2.* and *3.* □

5.1.3 The Coerciveness of J on $\mathcal{W}_0$

Let $\mathcal{A}_0 := \{\delta e = (\delta x, \delta y, \varepsilon_0, w) \in \mathcal{A} \colon \delta y = 0\}$. By [25, Lemma 6.3], $\delta e \in \mathcal{A}_0$ is orthogonal to $\mathcal{V}_0$ with respect to J, if and only if

$$\exists\, \omega_T \in V_T^{\perp} \text{ and } \theta_0 \in \mathbb{R} \text{ such that } J(\delta e, \overline{\delta e}) = \langle \omega_T\,,\, \overline{\delta x}\rangle + \theta_0 \overline{\varepsilon}_0 \quad \forall \overline{\delta e} \in \mathcal{A}_0.$$

If we choose $p(T) = -\omega_T$ in (25), taking into account that $\overline{\delta x} \in \mathbb{R}^n$, we obtain

$$\theta_0 = \varepsilon_0 J_0 + \langle p(\widehat{\tau}_2)\,,\, k(\widehat{x}_T)\rangle,$$
$$\langle p(t)\,,\, \dot{g}_t(\widehat{x}_T)\rangle + L_{\zeta(t)} L_{\dot{g}_t} \widetilde{\beta}\,(\widehat{x}_T) + w(t) R(t) \equiv 0. \tag{27}$$

By (27), the couple $(p(t), \zeta(t))$ solves (24) with $p(T) \in V_T^{\perp}$. Moreover

$$2\, J[\delta e]^2 = \varepsilon_0^2 J_0 + \varepsilon_0 \langle p(\widehat{\tau}_2)\,,\, k(\widehat{x}_T)\rangle \qquad \forall \delta e \in \mathcal{W}_0 \cap \mathcal{V}_0^{\perp_J}.$$

Using the properties of ι we can summarize this result in the following lemma:

Lemma 2 *Let J be coercive on $\mathcal{V}_0$. Then the followings are equivalent:*

1. *J is coercive on $\mathcal{W}_0$.*
2. *If $(\delta p, \delta x) \in L''_T$ and $\zeta''_{\widehat{\tau}_2}(\delta p, \delta x) = \varepsilon_0 k(\widehat{x}_T)$ with $(\delta x, \varepsilon_0) \neq (0,0)$, then*

$$\varepsilon_0^2 J_0 + \varepsilon_0 \langle \mu''_{\widehat{\tau}_2}(\delta p, \delta x)\,,\, k(\widehat{x}_T)\rangle = \varepsilon_0^2 J_0 + \langle \mu''_{\widehat{\tau}_2}(\delta p, \delta x)\,,\, \zeta''_{\widehat{\tau}_2}(\delta p, \delta x)\rangle > 0.$$

3. *If $(\delta\ell, \varepsilon_0) \in L_T \times \mathbb{R}$ and $(\pi\mathcal{H}_{\widehat{\tau}_2})_*\delta\ell = \varepsilon_0 \widetilde{k}(\widehat{x}_2)$ with $(\pi_*\delta\ell, \varepsilon_0) \neq (0,0)$, then*

$$\varepsilon_0^2 J_0 + \boldsymbol{\sigma}\left(\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell, (0, (\pi\mathcal{H}_{\widehat{\tau}_2})_*\delta\ell)\right) > 0.$$

5.1.4 The Coerciveness on $\mathcal{W}$

Again by [25, Lemma 6.3], a variation $\delta e \in \mathcal{A}$ is in $\mathcal{W}_0^{\perp_J}$ if and only if

$$\exists\, \omega_T \in V_T^\perp \text{ and } \omega_{\widehat{\tau}_2} \in (\mathbb{R}^n)^* \quad \text{s.t.} \quad J(\delta e, \overline{\delta e}) = \langle \omega_T\,,\, \overline{\delta x}\rangle + \langle \omega_{\widehat{\tau}_2}\,,\, \overline{\delta y}\rangle \quad \forall \overline{\delta e} \in \mathcal{A}.$$

Choosing $p(T) = -\omega_T$ in (25), taking into account that $\overline{\delta x} \in \mathbb{R}^n$ we obtain

$$\begin{gathered}\omega_{\widehat{\tau}_2} = \Gamma(\delta y, \cdot) + \varepsilon_0 L_{(\cdot)} L_k \widetilde{\beta}\,(\widehat{x}_T) + p(\widehat{\tau}_2),\\ \varepsilon_0 J_0 + L_{\delta y} L_k \widetilde{\beta}\,(\widehat{x}_T) + \langle p(\widehat{\tau}_2)\,,\, k(\widehat{x}_T)\rangle = 0,\\ \langle p(t)\,,\, \dot{g}_t(\widehat{x}_T)\rangle + L_{\zeta(t)} L_{\dot{g}_t} \widetilde{\beta}\,(\widehat{x}_T) + w(t)R(t) \equiv 0.\end{gathered}$$

Thus, $(p(t), \zeta(t))$ solves (24) with $p(T) \in V_T^\perp$ and

$$2\,J[\delta e]^2 = \Gamma[\delta y]^2 + \varepsilon_0 L_{\delta y} L_k \widetilde{\beta}\,(\widehat{x}_T) + \langle p(\widehat{\tau}_2)\,,\, \delta y\rangle, \quad \forall \delta e \in \mathcal{W} \cap \mathcal{W}_0^{\perp_J}. \tag{28}$$

We have thus obtained the following lemma:

Lemma 3 *Assume J is coercive on $\mathcal{W}_0$. Then the followings are equivalent.*

1. *J is coercive on $\mathcal{W}$.*
2. *If $(\delta p, \delta x) \in L''_T$ and $\zeta''_{\widehat{\tau}_2}(\delta p, \delta x) = \delta y + \varepsilon_0 k(\widehat{x}_T)$ with $(\delta x, \delta y, \varepsilon_0) \neq 0$ and $\varepsilon_0 H_{12}(\widehat{\ell}_1) + L_{\zeta''_{\widehat{\tau}_2}(\delta p, \delta x)} L_k \widetilde{\beta}\,(\widehat{x}_T) + \langle \mu''_{\widehat{\tau}_2}(\delta p, \delta x)\,,\, k(\widehat{x}_T)\rangle = 0$, then*

$$\begin{aligned}\Gamma[\zeta''_{\widehat{\tau}_2}(\delta p, \delta x)]^2 &- 2\varepsilon_0 L_{\zeta''_{\widehat{\tau}_2}(\delta p, \delta x)} L_k \widehat{\alpha}\,(\widehat{x}_T) + \varepsilon_0^2 L_k^2 \widehat{\alpha}\,(\widehat{x}_T)\\ &+ \langle \mu''_{\widehat{\tau}_2}(\delta p, \delta x)\,,\, \zeta''_{\widehat{\tau}_2}(\delta p, \delta x)\rangle + \varepsilon_0^2 H_{12}(\widehat{\ell}_1) > 0.\end{aligned} \tag{29}$$

3. *If $(\delta\ell, \varepsilon_0, \delta y) \in L_T \times \mathbb{R} \times T_{\widehat{x}_T}\widehat{N}_0$, and $(\pi\mathcal{H}_{\widehat{\tau}_2})_*\delta\ell = \widehat{S}_{\widehat{\tau}_2 *}\delta y + \varepsilon_0 \widetilde{k}(\widehat{x}_2)$ with $(\pi_*\delta\ell, \varepsilon_0, \delta y) \neq (0,0,0)$, and $\varepsilon_0 H_{12}(\widehat{\ell}_1) - \boldsymbol{\sigma}\left(\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell, \overrightarrow{\widetilde{K}}(\widehat{\ell}_2)\right) = 0$, then*

$$\boldsymbol{\sigma}\left(\mathrm{d}(\widehat{\alpha} \circ \widehat{S}_{\widehat{\tau}_2}^{-1})_* \widehat{S}_{\widehat{\tau}_2 *}\delta y, -\varepsilon_0 \overrightarrow{\widetilde{K}}(\widehat{\ell}_2) + \mathcal{H}_{\widehat{\tau}_2 *}\delta\ell\right) > 0. \tag{30}$$

Proof By the antisymplectic isomorphism ι the conditions required either in *2.* or in *3.* are equivalent and they mean that $\delta e = (\delta x, \delta y, \varepsilon_0, w) \in \mathcal{W}_0^{\perp_J}$. It is easy to see that the left-hand side of (29) coincides with (28). We have to prove that inequality (29) is equivalent to (30). By the conditions in *2.* we get

$$2J[\delta e]^2 = \Gamma[\zeta''_{\widehat{\tau}_2}(\delta p, \delta x)]^2 - 2\varepsilon_0 L_{\zeta''_{\widehat{\tau}_2}(\delta p, \delta x)} L_k \widehat{\alpha}\,(\widehat{x}_T) + \varepsilon_0^2 L_k^2 \widehat{\alpha}\,(\widehat{x}_T)$$

$$
\begin{aligned}
&+ \langle \mu''_{\widehat{\tau}_2}(\delta p,\delta x)\,,\,\zeta''_{\widehat{\tau}_2}(\delta p,\delta x)\rangle + \varepsilon_0^2 H_{12}(\widehat{\ell}_1)\\
&= \Gamma[\zeta''_{\widehat{\tau}_2}(\delta p,\delta x)]^2 - 2\varepsilon_0 L_{\zeta''_{\widehat{\tau}_2}(\delta p,\delta x)} L_k \widehat{\alpha}\,(\widehat{x}_T) + \varepsilon_0^2 L_k^2 \widehat{\alpha}\,(\widehat{x}_T)\\
&\quad + \langle \mu''_{\widehat{\tau}_2}(\delta p,\delta x)\,,\,\delta y\rangle - \varepsilon_0 L_{\zeta''_{\widehat{\tau}_2}(\delta p,\delta x)} L_k \widetilde{\beta}\,(\widehat{x}_T)\\
&= \Gamma(\zeta''_{\widehat{\tau}_2}(\delta p,\delta x),\delta y) - \varepsilon_0 L_{\delta y} L_k \widehat{\alpha}\,(\widehat{x}_T) + \langle \mu''_{\widehat{\tau}_2}(\delta p,\delta x)\,,\,\delta y\rangle\\
&= \boldsymbol{\sigma}\left(\begin{pmatrix}\mu''_{\widehat{\tau}_2}(\delta p,\delta x)\\ \zeta''_{\widehat{\tau}_2}(\delta p,\delta x)\end{pmatrix} - \begin{pmatrix}\mu''_{\widehat{\tau}_2}(\delta p,\delta x)\\ 0\end{pmatrix}, \mathrm{d}(-\widehat{\alpha}-\widetilde{\beta})_*\delta y\right)\\
&\quad - \varepsilon_0 L_{\delta y} L_k \widehat{\alpha}\,(\widehat{x}_T) + \langle \mu''_{\widehat{\tau}_2}(\delta p,\delta x)\,,\,\delta y\rangle\\
&= \boldsymbol{\sigma}\left(\mathcal{H}''_{\widehat{\tau}_2}(\delta p,\delta x), \mathrm{d}(-\widehat{\alpha}-\widetilde{\beta})_*\delta y\right) - \varepsilon_0 L_{\delta y} L_k \widehat{\alpha}\,(\widehat{x}_T)\,.
\end{aligned}
$$

Applying ι and setting $\delta\ell := \iota^{-1}(\delta p,\delta x)$ we get

$$
\begin{aligned}
2J[\delta e]^2 &= \boldsymbol{\sigma}\left(\mathrm{d}\widehat{\alpha}_*\delta y, \widehat{\mathcal{F}}^{-1}_{\widehat{\tau}_2 *}\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell\right) - \varepsilon_0\boldsymbol{\sigma}\left(\mathrm{d}\widehat{\alpha}_*\delta y, \overrightarrow{K}(\widehat{\ell}_T)\right)\\
&= \boldsymbol{\sigma}\left(\mathrm{d}\widehat{\alpha}_*\delta y, \widehat{\mathcal{F}}^{-1}_{\widehat{\tau}_2 *}\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell - \varepsilon_0\overrightarrow{K}(\widehat{\ell}_T)\right).
\end{aligned}
$$
□

5.2 The Free Final Point Case

If $N_T = \mathbb{R}^n$, we fall into case 2, the extremal $\widehat{\lambda}$ is normal ($p_0 = 1$) and $\beta = c_T$. The coerciveness of J_{ext} amounts to $L^2_{f_\mathrm{d}} c_T\,(\widehat{x}_T) > 0$ together with the coerciveness of J on $\mathcal{W}$ where $\widetilde{c} = \widetilde{\beta}$ and $V_T = \mathbb{R}^n$.

In this case $L_T = \left\{\mathrm{d}(-\widetilde{c})_*\delta x \colon \delta x \in \mathbb{R}^n\right\}$, the coerciveness on $\mathcal{W}_0$ is studied in [1] and the coerciveness on $\mathcal{W} \cap \mathcal{W}_0^{\perp_J}$, i.e. Lemma 3, reads

Lemma 4 *The quadratic form J is coercive on $\mathcal{W}$ if and only if it is coercive on $\mathcal{W}_0$ and for any non trivial triple $(\delta x, \varepsilon_0, \delta y) \in \mathbb{R}^n \times \mathbb{R} \times T_{\widehat{x}_T}\widehat{N}_0$ such that*

$$
(\pi\mathcal{H}_{\widehat{\tau}_2})_*\,\mathrm{d}(-\widetilde{c})_*\delta x = \widehat{S}_{\widehat{\tau}_2 *}\delta y + \varepsilon_0\widetilde{k}(\widehat{x}_2), \tag{31}
$$

$$
\varepsilon_0 H_{12}(\widehat{\ell}_1) - \boldsymbol{\sigma}\left(\mathcal{H}_{\widehat{\tau}_2 *}\,\mathrm{d}(-\widetilde{c})_*\delta x, \overrightarrow{\widetilde{K}}(\widehat{\ell}_2)\right) = 0, \tag{32}
$$

there holds

$$
\boldsymbol{\sigma}\left(\mathrm{d}(\widehat{\alpha}\circ\widehat{S}^{-1}_{\widehat{\tau}_2})_*\widehat{S}_{\widehat{\tau}_2 *}\delta y, -\varepsilon_0\overrightarrow{\widetilde{K}}(\widehat{\ell}_2) + \mathcal{H}_{\widehat{\tau}_2 *}\,\mathrm{d}(-\widetilde{c})_*\delta x\right) > 0. \tag{33}
$$

5.3 Reduction to the Free Final Point Case

We now prove how, thanks to [25, Theorem 13.2], we can add a penalty to problem (20)–(21) and reduce ourselves to the free final point case $N_T = \mathbb{R}^n$.

Let r be the dimension of the linear space V_T and let $f_2, \ldots, f_r$ be vector fields tangent to N_T. In case 1. choose $f_1 = f_\mathrm{d}$, while in case 2. let f_1 be a vector field tangent to N_T and such that $f_1(\widehat{x}_T) = f_\mathrm{d}(\widehat{x}_T)$. In both cases we get $\mathrm{Span}\,\{f_1(\widehat{x}_T), f_2(\widehat{x}_T), \ldots f_r(\widehat{x}_T)\} = V_T$. Choose coordinates as follows

$$
x \mapsto \exp x^1 f_1 \circ x^2 f_2 \ldots \circ \exp x^r f_r \circ \exp x^{r+1} f_{r+1} \circ \ldots \circ \exp x^n f_n(\widehat{x}_T)
$$

and let $\psi\colon \mathbb{R}^n \to \mathbb{R}$ be defined as $\psi(x) := \sum_{i=r+1}^{n}(x^i)^2$.

Consider the Hilbert space $\mathcal{B} := \left\{\delta e = (\delta x, \delta y, \varepsilon_0, w) \in \mathcal{A}\colon \delta y \in T_{\widehat{x}_T}\widehat{N}_0\right\}$ and apply [25, Theorem 13.2] to the Legendre form J given by (20) and to the weakly-continuous form $\mathrm{D}^2\psi(\widehat{x}_T)[\delta e]^2 := \mathrm{D}^2\psi(\widehat{x}_T)[\delta x]^2$. By such theorem there exists $\rho > 0$ such that

$$J_\rho[\delta e]^2 := J[\delta e]^2 + \frac{\rho}{2}\,\mathrm{D}^2\psi(\widehat{x}_T)[\delta e]^2$$

is coercive on $\mathcal{B}$. Moreover, defining

$$\widetilde{c} := \widetilde{\beta} + \rho\psi, \qquad \widetilde{\Gamma} := \mathrm{D}^2(\widehat{\alpha} + \widetilde{c})(\widehat{x}_T), \tag{34}$$

we can prove the following result.

Lemma 5 *For any $\delta e \in \mathcal{B}$ the quadratic form J_ρ can be written as*

$$\begin{aligned} J_\rho[\delta e]^2 = {} & \frac{1}{2}\widetilde{\Gamma}[\delta y]^2 + \frac{\varepsilon_0^2}{2}\left(L_k^2\widetilde{c}\,(\widehat{x}_T) + H_{12}(\widehat{\ell}_1)\right) \\ & + \varepsilon_0 L_{\delta y}L_k\widetilde{c}\,(\widehat{x}_T) + \frac{1}{2}\int_{\widehat{\tau}_2}^{T}\left(2\,w(t)L_{\zeta(t)}L_{\dot{g}_t}\widetilde{c}\,(\widehat{x}_T) + w(t)^2R(t)\right)\mathrm{d}t. \end{aligned} \tag{35}$$

Proof Let $\Psi[\delta e]^2$ be the right-hand side of (35). For $\delta e \in \mathcal{B}$ we have

$$\begin{aligned} 2J_\rho[\delta e]^2 &= \mathrm{D}^2(\widehat{\alpha} + \widetilde{\beta} \pm \rho\psi)(\widehat{x}_T)[\delta y]^2 + \varepsilon_0^2 L_k^2(\widetilde{\beta} \pm \rho\psi)\,(\widehat{x}_T) \\ &\quad + 2\varepsilon_0 L_{\delta y}L_k(\widetilde{\beta} \pm \rho\psi)\,(\widehat{x}_T) + \varepsilon_0^2 H_{12}(\widehat{\ell}_1) + \mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta x]^2 \\ &\quad + \int_{\widehat{\tau}_2}^{T}\left(2\,w(t)L_{\zeta(t)}L_{\dot{g}_t}(\widetilde{\beta} \pm \rho\psi)\,(\widehat{x}_T) + w(t)^2R(t)\right)\mathrm{d}t \\ &= 2\Psi[\delta e]^2 + \mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta x]^2 - \Big(\mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta y]^2 + \varepsilon_0^2 L_k^2(\rho\psi)\,(\widehat{x}_T) \\ &\quad + 2\varepsilon_0 L_{\delta y}L_k(\rho\psi)\,(\widehat{x}_T) + \int_{\widehat{\tau}_2}^{T} 2\,w(t)L_{\zeta(t)}L_{\dot{g}_t}(\rho\psi)\,(\widehat{x}_T)\,\mathrm{d}t\Big) \\ &= 2\Psi[\delta e]^2 + \mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta x]^2 - \mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta y + \varepsilon_0 k]^2 \\ &\quad - \int_{\widehat{\tau}_2}^{T} 2\,w(t)L_{\zeta(t)}L_{\dot{g}_t}(\rho\psi)\,(\widehat{x}_T)\,\mathrm{d}t. \end{aligned}$$

Recalling that $\mathrm{d}\psi(\widehat{x}_T) = 0$ and $\dot{\zeta}(t) = w(t)\dot{g}_t(\widehat{x}_T)$, we observe that

$$\begin{aligned} \int_{\widehat{\tau}_2}^{T} 2\,w(t)L_{\zeta(t)}L_{\dot{g}_t}(\rho\psi)\,(\widehat{x}_T)\,\mathrm{d}t &= \int_{\widehat{\tau}_2}^{T} 2\,\mathrm{D}^2(\rho\psi)(\widehat{x}_T)\left(\zeta(t), \dot{\zeta}(t)\right)\mathrm{d}t \\ &= \mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta x]^2 - \mathrm{D}^2(\rho\psi)(\widehat{x}_T)[\delta y + \varepsilon_0 k]^2. \end{aligned}$$

This completes the proof. □

Remark 6 The function $\widetilde{c}$ satisfies the properties required in equation (6) and $\widehat{\lambda}$ is a normal Pontryagin extremal for the optimal control problem (7a)–(7b). We recall that $\widetilde{c} = \widetilde{\beta}$ on N_T, so that $\widetilde{c}(x) \le p_0 c_T(x)$ for any $x \in N_T$.

6 The Main Results

We can now state the main theorems of this paper. The first theorem covers the case when the extended second variation is coercive. In Section 6.1 we give an analogous result in a special case when coerciveness is lost.

Theorem 2 *Assume that the reference trajectory $\widehat{\xi}$ satisfies Assumptions 1-5 and that the extended second variation $J_{\rm ext}$ is coercive. Then, if $p_0 = 1$, $\widehat{\xi}$ is a strict strong locally optimal trajectory of* (1a)–(1b)*; if $p_0 = 0$, $\widehat{\xi}$ is an isolated admissible trajectory of* (1b) *with respect to the C^0 distance between trajectories.*

Proof In this proof we refer to problem (7a)–(7b) choosing $\widetilde{c}$ as the function defined in (34) and α as the function defined in (14). The coerciveness of J_ρ on $\mathcal{B}$ is equivalent to the coerciveness of J_ρ on $\mathcal{B}_0 := \left\{\delta e \in \mathcal{B} \colon \delta y = 0\right\}$ and the coerciveness of J_ρ on $\mathcal{B}_0^{\perp_{J_\rho}}$. By [1] the coerciveness on $\mathcal{B}_0$ implies that first assumption of Theorem 1 is satisfied. Let us now prove that also the second one holds.
For any $\delta\ell = {\rm d}(-\widetilde{c})_*\delta x \in L_T$, let $\widetilde{\delta\ell} := \mathcal{H}_{\widehat{\tau}_2 *}\delta\ell$, $\delta y := \widehat{S}_{0*}^{-1}(\pi\mathcal{H}_0)_*\delta\ell$. With this notation, equation (4) reads

$$H_{12}(\widehat{\ell}_1)\langle {\rm d}\tau_1(\widehat{\ell}_2)\,,\,\widetilde{\delta\ell}\rangle = \boldsymbol{\sigma}\left(\widehat{\mathcal{F}}_{\widehat{\tau}_2 *}^{-1}\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell, \overrightarrow{K}(\widehat{\ell}_T)\right),$$

so that, using (32) we obtain $\varepsilon_0 = \langle {\rm d}\tau_1(\widehat{\ell}_2)\,,\,\widetilde{\delta\ell}\rangle$. Substituting in (5) we obtain that also (31) is satisfied so that the variation $\left(\pi_*\delta\ell, \langle {\rm d}\tau_1(\widehat{\ell}_2)\,,\,\widetilde{\delta\ell}\rangle, \delta y\right)$ belongs to $\mathcal{B}\cap\mathcal{B}_0^{\perp_{J_\rho}}$. Thus, if $(\pi\mathcal{H}_0)_*\delta\ell \neq 0$, inequality (33) in Lemma 4 is satisfied. Notice that, by (31), inequality (33) can also be written as

$$\boldsymbol{\sigma}\left({\rm d}\left(\widehat{\alpha}\circ\widehat{S}_{\widehat{\tau}_2}^{-1}\right)_*\left(-\varepsilon_0\widetilde{k}(\widehat{x}_2) + \pi_*\widetilde{\delta\ell}\right), -\varepsilon_0\overrightarrow{\widetilde{K}}(\widehat{\ell}_2) + \widetilde{\delta\ell}\right) > 0. \tag{36}$$

Taking into account that (5) can equivalently be written as

$$\widehat{\mathcal{F}}_{\widehat{\tau}_2 *}\widehat{\mathcal{F}}_{0*}^{-1}\mathcal{H}_{0*}\delta\ell = \widehat{\mathcal{F}}_{\widehat{\tau}_2 *}\left(-\varepsilon_0\overrightarrow{K}(\widehat{\ell}_T) + \widehat{\mathcal{F}}_{\widehat{\tau}_2 *}^{-1}\mathcal{H}_{\widehat{\tau}_2 *}\delta\ell\right) = -\varepsilon_0\overrightarrow{\widetilde{K}}(\widehat{\ell}_2) + \widetilde{\delta\ell},$$

and substituting in (36) we finally get assumption *2.* of Theorem 1:

$$\begin{aligned} 0 &< \boldsymbol{\sigma}\left({\rm d}(\widehat{\alpha}\circ\widehat{S}_{\widehat{\tau}_2}^{-1})_*\left(\widehat{S}_{\widehat{\tau}_2 *}\widehat{S}_{0*}^{-1}(\pi\mathcal{H}_0)_*\delta\ell\right), \widehat{\mathcal{F}}_{\widehat{\tau}_2 *}\widehat{\mathcal{F}}_{0*}^{-1}\mathcal{H}_{0*}\delta\ell\right) \\ &= \boldsymbol{\sigma}\left({\rm d}\,\widehat{\alpha}_*\left(\widehat{S}_{0*}^{-1}(\pi\mathcal{H}_0)_*\delta\ell\right), \widehat{\mathcal{F}}_{0*}^{-1}\mathcal{H}_{0*}\delta\ell\right) = \boldsymbol{\sigma}\left({\rm d}\,\alpha_*(\pi\mathcal{H}_0)_*\delta\ell, \mathcal{H}_{0*}\delta\ell\right). \end{aligned}$$

Thus $\widehat{\xi}$ is a strict strong local minimizer for problem (7a)–(7b) i.e. if $\xi \neq \widehat{\xi}$ is an admissible trajectory for problem (7a)–(7b), then $\widetilde{C}(\widehat{\xi}) < \widetilde{C}(\xi)$. Let ξ be an admissible trajectory for problem (1a)–(1b) whose graph is sufficiently close to the graph of $\widehat{\xi}$. Recalling the definitions of C and $\widetilde{C}$ and Remark 6 we obtain

$$p_0 C(\xi) \geq \alpha(\xi(0)) + \widetilde{c}(\xi(T)) \geq \alpha(\widehat{x}_0) + \widetilde{c}(\widehat{x}_T) = p_0 C(\widehat{\xi}) \tag{37}$$

and equality holds if and only if $\xi \equiv \widehat{\xi}$. This completes the proof in the normal case $p_0 = 1$. In the abnormal case $p_0 = 0$, equation (37) yields $\widetilde{C}(\xi) = \widetilde{C}(\widehat{\xi})$. Hence $\xi = \widehat{\xi}$, i.e. $\widehat{\xi}$ is an isolated admissible trajectory. □

6.1 A Non-Coercive Case

We now consider the case when the singular vector field $f_{\mathrm{d}} := h_3 - h_2$ is not null at the final point $\widehat{x}_T$ of the reference trajectory, and there is a neighbourhood $\mathcal{U}_T$ of $\widehat{x}_T$ such that

$$f_{\mathrm{d}}(x) \in T_x N_T, \qquad L_{f_{\mathrm{d}}}(p_0 c_T)(x) = 0 \qquad \forall x \in \mathcal{U}_T \cap N_T. \tag{38}$$

In this case the coerciveness of J_{ext} fails, since $\delta e = (f_{\mathrm{d}}(\widehat{x}_T), 0, 0, -1, 0)$ is an admissible variation and $J_{\mathrm{ext}}[\delta e]^2 = 0$. By the properties of f_{d}, we can choose the extension β such that $L_{f_{\mathrm{d}}}\beta \equiv 0$ and β satisfies the properties of $\widetilde{\beta}$ in (17). If we set $\widetilde{\beta} := \beta$ and $V_T := T_{\widehat{x}_T} N_T$, we obtain that the coerciveness of J_{ext} on the subspace of $\mathcal{W}_{\mathrm{ext}}$ such that $\varepsilon_1 = 0$ is the coerciveness of J given in (20) on $\mathcal{W}$ defined in (21).

Theorem 3 *Let the reference trajectory $\widehat{\xi}$ satisfy Assumptions 1–5 and relations* (38) *hold. Assume $f_{\mathrm{d}}(\widehat{x}_T) \neq 0$ and J_{ext} restricted to $\{\delta e \in \mathcal{W}_{\mathrm{ext}} \colon \varepsilon_1 = 0\}$ is coercive. If $p_0 = 1$, then $\widehat{\xi}$ is a strict strong locally optimal trajectory of* (1a)–(1b). *If $p_0 = 0$, then $\widehat{\xi}$ is an isolated admissible trajectory of* (1b) *with respect to the C^0 distance between trajectories.*

Proof In this case the reduction of the problem to the free final point case considered in Section 5.3 can be done by choosing $f_1 \equiv f_{\mathrm{d}}$. Therefore, the function $\widetilde{c}$ is now given by $\widetilde{c} = \beta + \rho\psi$, so that $\widetilde{c} = p_0 c_T$ on N_T. With this new $\widetilde{c}$ and J_ρ defined accordingly, we can exploit the coerciveness of J_ρ on $\mathcal{B}$ just as in the proof of Theorem 2 and thus prove our claim. $\square$

7 Examples

A particularly meaningful group of examples are those given by Bolza problems. It is well known that Bolza problems can be transformed into Mayer ones by adding a x^0-coordinate for the functional. Here we consider a particular class of problems which fit into the non-coercive case studied in Section 6.1 and which are present in the literature; see [16, 26]. Indeed we consider the problem

$$\text{minimize} \;\; C(\xi) = c_0(\xi(0)) + \int_0^T c(\xi(t))\, \mathrm{d}t \text{ subject to} \tag{39}$$

$$\dot{\xi}(t) \in \mathcal{X}(\xi(t)), \quad \text{a.e. } t \in [0, T], \quad \xi(0) \in N_0, \quad \xi(T) \in N_T \tag{40}$$

i.e. a Bolza problem with no cost on the final point and where the running cost depends only on the state.

Assume $\widehat{\xi}$ is a bang-bang-singular trajectory which satisfies PMP for problem (39)–(40). As usual in the literature, we transform the problem into a Mayer problem in $\mathbb{R}^{n+1}$ by adding the x^0-coordinate, with the final cost given by $c_T(x^0, x) = x^0$. If f_{d} is not null and tangent to N_T in a neighbourhood of $\widehat{x}_T$, then we fall in the case of Theorem 3. In particular, we recall that, when $N_T = \mathbb{R}^n$, then SGLC yields $f_{\mathrm{d}}(\widehat{x}_T) \neq 0$; see [1].

In [26] the author shows the existence of a bang-bang-singular extremal for a Van der Pool oscillator. In [20] such extremal is proved to be optimal among extremals having the same bang-bang-singular structure. In [1] the authors, applying the theory shown here, prove that such extremal is indeed a strong local minimizer.

Another problem fitting in this class can be found in [16], where the dynamical constraint is given by the Rayleigh equation in $\mathbb{R}^2$. Indeed, this problem can be written, following our notation, as (39)-(40), setting

$$T = 4.5, \qquad c_0(x) \equiv 0, \qquad c(x) := x_1^2 + x_2^2,$$
$$N_0 = \left\{(-5,-5)^t\right\}, \qquad N_T = \mathbb{R}^2,$$
$$\mathcal{X}(x) = \Big(x_2\ ,\ -2x_1 + x_2(1.4 - 0.14x_2^2) + k\Big)^t, \qquad k \in [k_1, k_2],$$

where k_1, k_2 are given real numbers such that $k_1 < k_2$. The authors show that in the case $k_1 = -8$, $k_2 = 0$ there exists a Pontryagin extremal which has a bang-bang-singular structure and that the associated trajectory $\widehat{\xi}$ is optimal among trajectories associated to the same structure.

Up to the authors knowledge only examples with fixed-free end point constraints are in the literature. Nevertheless, if N_T is any manifold in $\mathbb{R}^2$ such that $\widehat{\xi}(T) \in N_T$, e.g. $N_T = \left\{x \in \mathbb{R}^2 \colon x_1 = \widehat{\xi}_1(T)\right\}$, then $\widehat{\xi}$ is a Pontryagin extremal also for the final end point constraint $\xi(T) \in N_T$. Anyway a numerical investigation, which goes behind the scopes of this paper, is in order.

8 Conclusions

In this paper, we provide sufficient second order conditions for strong local optimality of bang-bang-singular extremals. Usually, these concatenations are studied for single input control systems. The paper deals with multi input systems, which are reduced to single input ones by taking advantage of the Hamiltonian methods. This reduction is allowed by the regularity conditions.

Perspective work includes the extension of the present results to bang-singular-bang extremals and the study of structural stability for the bang-bang-singular concatenations studied here. Indeed, Hamiltonian methods have proven to be a valid instrument for studying structural stability of strong local minimizers; see [27–30].

Acknowledgements The first author acknowledges the partial support given by INDAM-GNAMPA.

Appendices

A Computation of the Extended Second Variation

Problem (15) can be written as

$$\text{minimize } \Delta C(y, \delta v_0, \delta v) := (\widehat{\alpha}(y) + \beta(\eta(T))) - (\widehat{\alpha}(\widehat{x}_T) + \beta(\widehat{x}_T)) \quad \text{subject to}$$

$$\dot{\eta}(t) = \varphi_t(\eta(t)) := \begin{cases} \delta v_0(t) k_1(\eta(t)), & t \in]0, \widehat{\tau}_1[, \\ \delta v_0(t) k_2(\eta(t)), & t \in]\widehat{\tau}_1, \widehat{\tau}_2[, \\ \delta v(t) g_t(\eta(t)), & t \in]\widehat{\tau}_2, T[, \end{cases} \tag{41}$$

$$\delta v_0(t) > -1, \ \int_0^{\widehat{\tau}_2} \delta v_0(t) \, \mathrm{d}t = 0, \quad |\delta v(t)| < \delta, \quad \eta(0) = y \in \widehat{N}_0, \quad \eta(T) = x \in N_T.$$

We can allow for the controls δv_0, δv to be in $L^2 := L^2([0,T],\mathbb{R})$ since equation (41) is linear with respect to the controls. Defining

$$\gamma \colon y \in \mathbb{R}^n \mapsto (\widehat{\alpha} + \beta)(y) - (\widehat{\alpha} + \beta)(\widehat{x}_T) \in \mathbb{R},$$
$$\mathcal{L} \colon (t, y, \delta v_0, \delta v) \in [0,T] \times \mathbb{R}^n \times L^2 \times L^2 \mapsto \mathcal{L}_t(y, \delta v_0, \delta v) := \langle \mathrm{d}\beta(y) \, , \, \varphi_t(y) \rangle \in \mathbb{R}$$

we get

$$\Delta C(y, \delta v_0, \delta v) := \gamma(y) + \int_0^T \mathcal{L}_t(\eta(t), \delta v_0(t), \delta v(t)) \, \mathrm{d}t.$$

We aim at computing the second order approximation C'' of ΔC.
By the properties of γ, the constraint on δv_0, and PMP it is not difficult to see that

$$\partial_y \Delta C(\widehat{x}_T, 0, 0) = \mathrm{d}\gamma(\widehat{x}_T) = 0,$$
$$\langle \partial_{\delta v} \Delta C(\widehat{x}_T, 0, 0) \, , \, \delta v \rangle = \int_{\widehat{\tau}_2}^T \delta v(t) L_{g_t} \beta \, (\widehat{x}_T) \, \mathrm{d}t = -\int_{\widehat{\tau}_2}^T \delta v(t) F_{\mathrm{d}}(\widehat{\lambda}(t)) \, \mathrm{d}t = 0,$$
$$\langle \partial_{\delta v_0} \Delta C(\widehat{x}_T, 0, 0) \, , \, \delta v_0 \rangle = \int_0^{\widehat{\tau}_1} \delta v_0(t) L_{k_1} \beta \, (\widehat{x}_T) \, \mathrm{d}t + \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v_0(t) L_{k_2} \beta \, (\widehat{x}_T) \, \mathrm{d}t$$
$$= -H_1(\widehat{\ell}_1) \int_0^{\widehat{\tau}_1} \delta v_0(t) \, \mathrm{d}t - H_2(\widehat{\ell}_1) \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v_0(t) \, \mathrm{d}t = -H_1(\widehat{\ell}_1) \int_0^{\widehat{\tau}_2} \delta v_0(t) \, \mathrm{d}t = 0.$$

Thus, the first order approximation is null and the second order approximation is intrinsecally well defined. Obviously,

$$\partial^2_{yy} C(\widehat{x}_T, 0, 0) = \mathrm{D}^2 \gamma(\widehat{x}_T).$$

Denote as $\mathcal{L}''_t$ the second order derivative of $\mathcal{L}_t$ at $(\widehat{x}_T, 0, 0)$ and let $\delta\eta$ be the linearization of η, i.e. $\delta\eta$ solves the problem

$$\dot{\delta\eta}(t) = \varphi_t(\widehat{x}_T), \quad \delta\eta(0) = \delta y \in T_{\widehat{x}_T} \widehat{N}_0, \quad \delta\eta(T) = \delta x \in T_{\widehat{x}_T} N_T. \tag{42}$$

Hence

$$2\, C''[\delta y, \delta v_0, \delta v]^2 = \mathrm{D}^2 \gamma(\widehat{x}_T)[\delta y]^2 + \int_0^T \mathcal{L}''_t[\delta\eta(t), \delta v_0(t), \delta v(t)]^2 \, \mathrm{d}t.$$

For the sake of computations, let us define

$$I_1(\delta y, \delta v_0) := \int_0^{\widehat{\tau}_1} \mathcal{L}''_t[\delta\eta(t), \delta v_0(t), 0]^2 \, \mathrm{d}t = 2 \int_0^{\widehat{\tau}_1} \delta v_0(t) L_{\delta\eta(t)} L_{k_1} \beta \, (\widehat{x}_T) \, \mathrm{d}t,$$
$$I_2(\delta y, \delta v_0) := \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \mathcal{L}''_t[\delta\eta(t), \delta v_0(t), 0]^2 \, \mathrm{d}t = 2 \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v_0(t) L_{\delta\eta(t)} L_{k_2} \beta \, (\widehat{x}_T) \, \mathrm{d}t,$$
$$I_3(\delta y, \delta v) := \int_{\widehat{\tau}_2}^T \mathcal{L}''_t[\delta\eta(t), 0, \delta v(t)]^2 \, \mathrm{d}t = 2 \int_{\widehat{\tau}_2}^T \delta v(t) L_{\delta\eta(t)} L_{g_t} \beta \, (\widehat{x}_T) \, \mathrm{d}t.$$

Let $\varepsilon_0 := \int_0^{\widehat{\tau}_1} \delta v_0(s) \, \mathrm{d}s$. Then

$$\delta\eta(t) = \delta y + \int_0^t \delta v_0(s) \, \mathrm{d}s \, k_1(\widehat{x}_T), \qquad t \in [0, \widehat{\tau}_1],$$
$$\delta\eta(t) = \delta y + \varepsilon_0 k_1(\widehat{x}_T) + \int_{\widehat{\tau}_1}^t \delta v_0(s) \, \mathrm{d}s \, k_2(\widehat{x}_T), \qquad t \in [\widehat{\tau}_1, \widehat{\tau}_2],$$

$$\delta\eta(t) = \delta y + \varepsilon_0 k(\widehat{x}_T) + \int_{\widehat{\tau}_2}^{t} \delta\upsilon(s) g_s(\widehat{x}_T)\,\mathrm{d}s\,, \qquad t \in [\widehat{\tau}_2, T].$$

In particular

$$\begin{aligned} I_1(\delta y, \delta\upsilon_0) &= 2\varepsilon_0 L_{\delta y} L_{k_1}\beta\,(\widehat{x}_T)\,\mathrm{d}t + \varepsilon_0^2 L_{k_1}^2\beta\,(\widehat{x}_T)\,,\\ I_2(\delta y, \delta\upsilon_0) &= -2\varepsilon_0 L_{\delta y+\varepsilon_0 k_1} L_{k_2}\beta\,(\widehat{x}_T) + \varepsilon_0^2 L_{k_2}^2\beta\,(\widehat{x}_T)\,. \end{aligned}$$

Define $w(t) := \int_{\widehat{\tau}_2}^{t} -\delta\upsilon(s)\,\mathrm{d}s$, $\varepsilon_1 := w(T)$ and let $\zeta\colon [\widehat{\tau}_2, T] \to \mathbb{R}^n$ solve the Cauchy problem

$$\dot\zeta(t) = w(t)\dot g_t(\widehat{x}_T), \qquad \zeta(\widehat{\tau}_2) = \delta\eta(\widehat{\tau}_2).$$

By (42), $\zeta(T) = \delta x + \varepsilon_1 f_{\mathrm{d}}(\widehat{x}_T)$. Moreover, applying an intrinsic version of Goh transformation as in [9] we obtain

$$\begin{aligned} I_3(\delta y, \delta\upsilon) &= L_{\delta y+\varepsilon_0 k}\int_{\widehat{\tau}_2}^{T} -\dot w(t) L_{g_t}\beta\,(\widehat{x}_T)\,\mathrm{d}t + \int_{\widehat{\tau}_2}^{T} \dot w(t)\int_{\widehat{\tau}_2}^{t}\dot w(s) L_{g_s}L_{g_t}\beta\,(\widehat{x}_T)\,\mathrm{d}s\,\mathrm{d}t\\ &= -\varepsilon_1^2 L_{f_{\mathrm{d}}}^2\beta\,(\widehat{x}_T) - 2\varepsilon_1 L_{\delta x}L_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) + \int_{\widehat{\tau}_2}^{T}\left(2\,w(t)L_{\zeta(t)}L_{\dot g_t}\beta\,(\widehat{x}_T) + w(t)^2 R(t)\right)\,\mathrm{d}t. \end{aligned}$$

Thus

$$\begin{aligned} 2C''[\delta y, \delta\upsilon_0, \delta\upsilon]^2 &= \mathrm{D}^2\gamma(\widehat{x}_T)[\delta y]^2 + I_1 + I_2 + I_3\\ &= \mathrm{D}^2\gamma(\widehat{x}_T)[\delta y]^2 + 2\varepsilon_0 L_{\delta y}L_k\beta\,(\widehat{x}_T)\,\mathrm{d}t + \varepsilon_0^2\left(L_k^2\beta\,(\widehat{x}_T) + L_{[k_2,k_1]}\beta\,(\widehat{x}_T)\right)\\ &\quad - \varepsilon_1^2 L_{f_{\mathrm{d}}}^2\beta\,(\widehat{x}_T) - 2\varepsilon_1 L_{\delta x}L_{f_{\mathrm{d}}}\beta\,(\widehat{x}_T) + \int_{\widehat{\tau}_2}^{T}\left(2\,w(t)L_{\zeta(t)}L_{\dot g_t}\beta\,(\widehat{x}_T) + w(t)^2R(t)\right)\,\mathrm{d}t \end{aligned}$$

subject to

$$\dot\zeta(t) = w(t)\dot g_t(\widehat{x}_T), \quad \zeta(\widehat{\tau}_2) = \delta y + \varepsilon_0 k(\widehat{x}_T), \quad \delta x = \zeta(T) - \varepsilon_1 f_{\mathrm{d}}(\widehat{x}_T) \in T_{\widehat{x}_T}N_T.$$

Notice that $\delta\upsilon_0$ appears only through ε_0, while the immersion

$$\delta\upsilon \in L^2([\widehat{\tau}_2, T], \mathbb{R}) \mapsto (w(t), w(T)) \in L^2([\widehat{\tau}_2, T], \mathbb{R}) \times \mathbb{R}$$

is continuous and dense. Thus we can extend C'' to variations $\delta e := (\delta x, \delta y, \varepsilon_0, \varepsilon_1, w) \in \mathcal{W}_{\mathrm{ext}}$ as defined in Section 5, and the extension coincides with J_{ext}.

B Splitting of the Second Variation

Lemma 6 *Assume $f_{\mathrm{d}}(\widehat{x}_T) \in T_{\widehat{x}_T}N_T$. Then the coerciveness of J_{ext} on $\mathcal{W}_{\mathrm{ext}}$ splits into $L_{f_{\mathrm{d}}}^2\beta\,(\widehat{x}_T) > 0$ plus the coerciveness of J on $\mathcal{W}$.*

Proof We decompose $\delta x \in T_{\widehat{x}_T} N_T$ as $\delta x = \delta z + r f_{\mathrm{d}}(\widehat{x}_T)$, $\delta z \in T_{\widehat{x}_T} \widetilde{N}_T$, where $\widetilde{N}_T$ is the manifold defined in Section 5.1. We can compute

$$
\begin{aligned}
2\, J_{\mathrm{ext}}[\delta e]^2 &= \mathrm{D}^2(\widehat{\alpha} + \beta \pm \widetilde{\beta})(\widehat{x}_T)[\delta y]^2 + \varepsilon_0^2 \left(L_k^2(\beta \pm \widetilde{\beta})\,(\widehat{x}_T) + H_{12}(\widehat{\ell}_1) \right) \\
&\quad + 2\varepsilon_0 L_{\delta y} L_k (\beta \pm \widetilde{\beta})\,(\widehat{x}_T) - \varepsilon_1^2 L_{f_{\mathrm{d}}}^2 \beta\,(\widehat{x}_T) - 2\varepsilon_1 L_{\delta z + r f_{\mathrm{d}}} L_{f_{\mathrm{d}}} \beta\,(\widehat{x}_T) \\
&\quad + \int_{\widehat{\tau}_2}^{T} \left(2\, w(t) L_{\zeta(t)} L_{\dot{g}_t} (\beta \pm \widetilde{\beta})\,(\widehat{x}_T) + w(t)^2 R(t) \right) \mathrm{d}t \\
&= \Gamma[\delta y]^2 + \varepsilon_0^2 J_0 + 2\varepsilon_0 L_{\delta y} L_k \widetilde{\beta}\,(\widehat{x}_T) + \mathrm{D}^2(\beta - \widetilde{\beta})(\widehat{x}_T)[\delta y + \varepsilon_0 k]^2 \\
&\quad - \left(\varepsilon_1^2 + 2\varepsilon_1 r\right) L_{f_{\mathrm{d}}}^2 \beta\,(\widehat{x}_T) + \left. \mathrm{D}^2(\beta - \widetilde{\beta})(\widehat{x}_T)[\zeta(t)]^2 \right|_{t=\widehat{\tau}_2}^{t=T} \\
&\quad + \int_{\widehat{\tau}_2}^{T} \left(2\, w(t) L_{\zeta(t)} L_{\dot{g}_t} \widetilde{\beta}\,(\widehat{x}_T) + w(t)^2 R(t) \right) \mathrm{d}t \\
&= 2J[(\delta z + (r + \varepsilon_1) f_{\mathrm{d}}(\widehat{x}_T), \delta y, \varepsilon_0, w)]^2 \\
&\quad + (r + \varepsilon_1)^2\, \mathrm{D}^2(\beta - \widetilde{\beta})(\widehat{x}_T)[f_{\mathrm{d}}(\widehat{x}_T)]^2 - \left(\varepsilon_1^2 + 2\varepsilon_1 r\right) L_{f_{\mathrm{d}}}^2 \beta\,(\widehat{x}_T) \\
&= 2J[(\delta z + (r + \varepsilon_1) f_{\mathrm{d}}(\widehat{x}_T), \delta y, \varepsilon_0, w)]^2 + r^2 L_{f_{\mathrm{d}}}^2 \beta\,(\widehat{x}_T)\,.
\end{aligned}
$$

The above computation shows that the real variable r is decoupled and $\delta z + (r + \varepsilon_1) f_{\mathrm{d}}(\widehat{x}_T)$ is a generic vector $\delta x \in T_{\widehat{x}_T} N_T$. This proves the claim. □

References

1. Poggiolini, L., Stefani, G.: Strong local optimality for a bang-bang-singular extremal: The fixed-free case. SIAM Journal on Control and Optimization **56**(3), 2274–2294 (2018). DOI 10.1137/17M1140248. URL https://doi.org/10.1137/17M1140248
2. Poggiolini, L., Stefani, G.: Constrained bang-bang-singular extremals. In: CDC 2019 Proceedings. IEEE (2019)
3. Agrachev, A.A., Sachkov, Y.L.: Control Theory from the Geometric Viewpoint. Springer-Verlag (2004)
4. Agrachev, A.A., Stefani, G., Zezza, P.: Strong optimality for a bang-bang trajectory. SIAM J. Control Optimization **41**(4), 991–1014 (2002)
5. Poggiolini, L., Stefani, G.: State-local optimality of a bang-bang trajectory: a Hamiltonian approach. Systems & Control Letters **53**, 269–279 (2004)
6. Poggiolini, L., Spadini, M.: Strong local optimality for a bang-bang trajectory in a Mayer problem. SIAM Journal on Control and Optimization **49**(1), 140–161 (2011). DOI 10.1137/090771405. URL http://epubs.siam.org/doi/abs/10.1137/090771405
7. Poggiolini, L., Spadini, M.: Bang–bang trajectories with a double switching time in the minimum time problem. ESAIM: COCV (2015). DOI 10.1051/cocv/2015021
8. Stefani, G.: Strong optimality of singular trajectories. In: F. Ancona, A. Bressan, P. Cannarsa, F. Clarke, P.R. Wolenski (eds.) Geometric Control and Nonsmooth Analysis, *Series on Advances in Mathematics for Applied Sciences*, vol. 76, pp. 300–326. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ (2008). Pp. 361 ISBN: 978-981-277-606-8
9. Poggiolini, L., Stefani, G.: Bang-singular-bang extremals: sufficient optimality conditions. Journal of Dynamical and Control Systems **17**, 469–514 (2011). URL http://dx.doi.org/10.1007/s10883-011-9127-y. 10.1007/s10883-011-9127-y
10. Chittaro, F.C., Stefani, G.: Singular extremals in multi–input time–optimal problem: a sufficient condition. Control and Cybernetics **39**(4), 1029–1068 (2010)
11. Chittaro, F.C., Stefani, G.: Minimum-time strong optimality of a singular arc: The multi-input non involutive case. ESAIM: COCV **22**(3), 786–810 (2016). URL http://dx.doi.org/10.1051/cocv/2015026
12. Agrachev, A.A., Biolo, C.: Switching in time-optimal problem with control in a ball. SIAM Journal on Control and Optimization **56**(1), 183–200 (2018). DOI 10.1137/16M110304X. URL https://doi.org/10.1137/16M110304X

13. Agrachev, A.A., Biolo, C.: Optimality of broken extremals. Journal of Dynamical and Control Systems **25**(2), 289–307 (2019). DOI 10.1007/s10883-018-9416-9. URL https://doi.org/10.1007/s10883-018-9416-9
14. Caillau, J.B., Orieux, M.: Sufficient conditions for time optimality of systems with control on the disk. In: CDC 2019 Proceedings. IEEE (2019)
15. Caillau, J.B., Orieux, M.: Singularities of min time affine control systems. URL https://hal.inria.fr/hal-01718345
16. Maurer, H., Osmolovskii, N.: Second-order conditions for optimal control problems with mixed control-state constraints and control appearing linearly. In: 2013 IEEE 52ND ANNUAL CONFERENCE ON DECISION AND CONTROL (CDC), IEEE Conference on Decision and Control, pp. 514–519 (2013)
17. Aronna, M.S., Bonnans, J.F., Dmitruk, A.V., Lotito, P.A.: Quadratic order conditions for bang-singular extremals. NUMERICAL ALGEBRA, CONTROL AND OPTIMIZATION **2**(3), 511–546 (2012). Doi:10.3934/naco.2012.2.511
18. Bonnard, B., Caillau, J.B., Trélat, E.: Second order optimality conditions in the smooth case and applications in optimal control. ESAIM: COCV **13**(2), 207–236 (2007). DOI 10.1051/cocv:2007012. URL https://doi.org/10.1051/cocv:2007012
19. Hilscher, R.S., Zeidan, V.: Sufficiency and sensitivity for nonlinear optimal control problems on time scales via coercivity. ESAIM: COCV **24**(4), 1705–1734 (2018). DOI 10.1051/cocv/2017070. URL https://doi.org/10.1051/cocv/2017070
20. Vossen, G.: Switching time optimization for bang-bang and singular controls. Journal of Optimization Theory and Applications **144**(2), 409–429 (2010). DOI 10.1007/s10957-009-9594-4
21. Agrachev, A.A., Stefani, G., Zezza, P.: Symplectic methods for strong local optimality in the bang-bang case. In: J.G. A. Anzaldo-Meneses B. Bonnard, F. Monroy-Perez (eds.) Contemporary Trends in Non-linear Geometric Control Theory and its Applications. World Scientific, London (2002)
22. Stefani, G., Zezza, P.: Variational Methods in Imaging and Geometric Control, chap. A Hamiltonian approach to sufficiency in optimal control with minimal regularity conditions: Part I. De Gruyter (2016)
23. Agrachev, A.A., Stefani, G., Zezza, P.: An invariant second variation in optimal control. Internat. J. Control **71**(5), 689–715 (1998)
24. Stefani, G., Zezza, P.: Constrained regular LQ-control problems. SIAM J. Control Optim. **35**(3), 876–900 (1997)
25. Hestenes, M.R.: Applications of the theory of quadratic forms in Hilbert space to calculus of variations. Pacific J. Math **1**, 525–581 (1951)
26. Maurer, H.: Theory and applications of bang-bang and singular control problems. Conference slides (2007). 23rd IFIP TC 7 Conference on System Modelling and Optimization Kraków, Poland, July 23-27, 2007
27. Felgenhauer, U., Poggiolini, L., Stefani, G.: Optimality and stability result for bang–bang optimal controls with simple and double switch behaviour. CONTROL AND CYBERNETICS **38**(4B), 1305 – 1325 (2009)
28. Poggiolini, L., Stefani, G.: Structural stability for bang–singular–bang extremals in the minimum time problem. SIAM J. Control Optim. **51**(5), 3511–3531 (2013). URL http://dx.doi.org/10.1137/120895421
29. Poggiolini, L., Stefani, G.: A case study in strong optimality and structural stability of bang-singular extremals. In: G. Stefani, U. Boscain, J.P. Gauthier, A. Sarychev, M. Sigalotti (eds.) Geometric Control Theory and Sub-Riemannian Geometry, *Springer INdAM Series*, vol. 5, pp. 333–350. Springer International Publishing (2014). DOI 10.1007/978-3-319-02132-4_19
30. Poggiolini, L.: Structural stability of bang–bang trajectories with a double switching time in the minimum time problem. SIAM JOURNAL ON CONTROL AND OPTIMIZATION **55**(6), 3779–3798 (2017). DOI 10.1137/16M1083761